Remembering what could have happened: Neural correlates of episodic counterfactual thinking

Running Head: Remembering what could have happened

Remembering what could have happened: Neural correlates of episodic counterfactual thinking

(In press) *Neuropsychologia*

De Brigard, F[1,2 *], Addis, D.R.[3,4], Ford, J.H.[1], Schacter, D.L.[2], & Giovanello, K.S[1, 5].

[1] Department of Psychology – UNC, Chapel Hill

[2] Department of Psychology – Harvard University

[3] School of Psychology – The University of Auckland

[4] Centre for Brain Research – The University of Auckland

[5] Biomedical Research Imaging Center – UNC, Chapel Hill

* Corresponding author:

Felipe De Brigard
Postdoctoral Fellow
Department of Psychology
Harvard University
William James Hall, 884
Cambridge, MA 02138
Cell Phone: (919) 259 8114
Work: (617) 496 5909
E-mail: brigard@wjh.harvard.edu

Abstract

Recent evidence suggests that our capacities to remember the past and to imagine what might happen in the future largely depend on the same core brain network that includes the middle temporal lobe, the posterior cingulate/retrosplenial cortex, the inferior parietal lobe, the medial prefrontal cortex, and the lateral temporal cortex. However, the extent to which regions of this core brain network are also responsible for our capacity to think about what *could* have happened in our past, yet did not occur (i.e., episodic counterfactual thinking), is still unknown. The present study examined this issue. Using a variation of the experimental recombination paradigm (Addis et al., 2009), participants were asked both to remember personal past events and to envision alternative outcomes to such events while undergoing functional magnetic resonance imaging. Three sets of analyses were performed on the imaging data in order to investigate two related issues. First, a mean-centered spatiotemporal partial least square (PLS) analysis identified a pattern of brain activity across regions of the core network that was common to episodic memory and episodic counterfactual thinking. Second, a non-rotated PLS analysis identified two different patterns of brain activity for likely and unlikely episodic counterfactual thoughts, with the former showing significant overlap with the set of regions engaged during episodic recollection. Finally, a parametric modulation was conducted to explore the differential engagement of brain regions during counterfactual thinking, revealing that areas such as the parahippocampal gyrus and the right hippocampus were modulated by the subjective likelihood of counterfactual simulations. These results suggest that episodic counterfactual thinking engages regions that form the core brain network, and also that the subjective likelihood of our counterfactual thoughts modulates the engagement of different areas within this set of regions.

Keywords

Episodic memory, Counterfactual thinking, Imagination, Mental time travel, fMRI, Partial least squares.

1. Introduction

Traditionally, the notion of episodic memory has been used in reference to the psychological capacity to remember the past (Tulving, 1985). As a result, most research on episodic memory has focused on episodic recollection: the cognitive process of bringing past experiences back to mind (Tulving, 2002). However, recent evidence suggests striking commonalities between the cognitive and neural processes required to remember one's past and those required to imagine one's future (for recent reviews, see Schacter, Addis, & Buckner, 2008; Schacter, Addis, Hassabis, Martin, Spreng, & Szpunar, 2012; Szpunar, 2010). Evidence from three lines of research supports this claim. First, neuropsychological studies with populations known to have episodic memory deficits, such as patients with amnesia (Tulving, 1985; Klein, Loftus, & Kihlstrom, 2002; Hassabis, Kumaran, Vann, & Maguire, 2007; Race, Keane, & Verfaellie, 2011; but see Squire, van der Horst, McDuff, Frascino, Hopkins et al., 2010), severe depression (Dickson & Bates, 2005; Williams, Ellis, Tyers, Healy, Rose et al., 1996), schizophrenia (D'Argembeau, Raffard, & Van der Linden, 2008), Alzheimer's disease (Addis, Sacchetti, Ally, Budson, & Schacter, 2009) and mild cognitive impairment (Gamboz, De Vito, Brandimonte, Pappalardo, Galone et al., 2010) show that they also exhibit impairments when mentally simulating events that may happen in their future, a cognitive process that has come to be known as *episodic future thinking* (Atance & O'Neil, 2001; Szpunar, 2010). Similar parallels between remembering the past and imagining the future have been observed in young children (Atance & O'Neil, 2001; Suddendorf & Busby, 2005) as well as in healthy old adults (Addis, Wong, & Schacter, 2008; Addis, Musicaro, Pan, and Schacter, 2010; Gaesser, Sacchetti, Addis, & Schacter, 2011; Spreng & Levine,

2006). Second, behavioral studies examining the phenomenological characteristics of episodic memory and episodic future thinking indicate that both processes are supported by common cognitive mechanisms (D'Argembeau & Van der Linden, 2004; D'Argembeau, Stawarczyk, Majerus, Collette, Van del Linden et al., 2009; Szpunar & McDermott, 2008). Third, functional neuroimaging studies comparing episodic memory and future thinking have revealed a common "core" brain network that is engaged during both processes (Addis, Wong, Schacter, 2007; Addis & Schacter, 2008; Hassabis, Kumara, & Maguire, 2007; Szpunar, Watson, & McDermott, 2007; Okuda, Fujii, Ohtake, Tsukiura, Tanki et al., 2003). This core network, which overlaps substantially with the default network (Buckner, Andrews-Hanna, & Schacter, 2008), involves primarily the medial temporal lobes (including the hippocampus), the cingulate/retrosplenial cortex, the inferior parietal lobe, the medial prefrontal cortex, and the lateral temporal cortex (Buckner & Carroll, 2007; Schacter, Addis, & Buckner, 2007).

To account for the phenomenological, neural and cognitive commonalities between remembering one's past and imagining one's future, Schacter and Addis (2007) put forth the constructive episodic simulation hypothesis. According to this hypothesis, episodic future thinking relies on much the same neural mechanisms, and shares much of the same phenomenological characteristics, as episodic memory because both cognitive operations depend on similar processes. When we remember an event, episodic memory processes reintegrate representational contents from the encoded experience to reconstruct the unified mental simulation we call recollection. Similarly, when we engage in episodic future thinking, some of the same processes recombine components from past experiences into a novel, yet memory-dependent, simulation of what may occur in the

future. However, the finding of common activations during both processes is consistent with an alternative hypothesis: Thinking about the future need not involve the recombination of components, but rather, may entail the mere recasting of a previous experience as a future event. By this "recasting" account, thinking about the future would consist of a two-fold process: An initial recollection of a specific past experience, followed by imagining that experience occurring not in the past, but in the future. Thus, recasting could explain why brain regions related to episodic memory are engaged during episodic future thinking, without postulating the flexible recombination of episodic components, as suggested by the constructive episodic simulation hypothesis.

In a recent study, Addis and collaborators (Addis, Pan, Vu, Laiser, & Schacter, 2009) tested the constructive episodic simulation hypothesis as an alternative to the recasting view using an *experimental recombination procedure.* This paradigm consists of collecting episodic memories from participants in order to extract details from the reported episodes. Such event details or components are subsequently recombined during a scanning session in which they are employed as visual cues. Addis and colleagues presented participants with three components (i.e., person, object and place) extracted from participants' memories. In one condition, all components belonged to the same memory and participants were simply asked to remember the event to which such episodic details belonged. In a second condition, participants were presented with randomly recombined components of their memories and were asked to imagine a future event that would include such event details. Finally, in a third condition, participants were presented with randomly recombined components of their reported memories, but were asked to imagine an alternative *past* event including such disjoint event details.

Using spatiotemporal partial least squares analysis (PLS), Addis, Pan et al. (2009) found that all three conditions commonly activated regions of the core brain network. They interpreted this result as supporting the constructive simulation hypothesis, as opposed to the recasting account, insofar as the experimental procedure required episodic recombination of elements into imagined future and past events. Specifically, they suggest that this common activation may reflect the retrieval of episodic contents—a process that is necessary not only when remembering past events, but also when constructing imagined future or past events through a process of recombination.

Importantly, in addition to finding evidence in support of the overlap between remembering and imagining, Addis, Pan et al. (2009) found two distinguishable patterns of brain activity within this shared neural network. The spatiotemporal PLS analysis also identified one subsystem within the core brain network that was preferentially associated with the remembering task, and another subsystem preferentially associated with the future and past imagining tasks. However, Addis, Pan et al. (2009) did not examine an essential feature of simulations of what may happen in the future and what may have happened in the past: the subjective likelihood of those events. Namely, when prospecting, we usually simulate episodes of what we think is likely or probable to occur to us in the future (Weiler, Suchan, & Daum, 2010). Similarly, we entertain thoughts about alternative past events that we consider more or less likely to have happened. However, as Addis, Pan et al. (2009) point out in their discussion, by randomly recombining episodic details taken from multiple memories, participants may have been presented with possible, yet quite unlikely past events that otherwise would have never occurred to them. As such, it remains unclear what are the precise neural mechanisms

underlying our capacity to simulate alternative versions of specific past personal episodes that could have happened but did not actually occur—a cognitive process we call *episodic counterfactual thinking* (De Brigard & Giovanello, 2012).

It is worth noting that, although research on the cognitive neuroscience of counterfactual thinking—broadly defined as thoughts of what may have been (Roese, 1997; Byrne, 2002; Epstude & Roese, 2008)—is growing, most studies focus on the simulation of counterfactual alternatives to impersonal events and/or decision-making tasks confined to lab settings, and only a handful have employed stimuli extracted from the participant's own episodic autobiographical recollections. In one such study, De Brigard, Szpunar, and Schacter (in press) asked participants to remember negative, positive, and neutral autobiographical memories, and then simulate self-generated counterfactual alternatives to those memories once or four times. Repeated simulation decreased the perceived plausibility of the episodic counterfactual events. In a neuropsychological study, Beldarrain and colleagues (Beldarrain, Garcia-Monco, Astigarraga, Gonzalez, & Grafman, 2005), compared spontaneous versus non-spontaneous generation of episodic counterfactual thoughts in patients with prefrontal damage. Although previous studies of counterfactual thinking have shown that the medial prefrontal and orbitofrontal cortices are critical for counterfactual thinking in decision making tasks (Barbey, Krueger, & Grafman, 2009), Beldarrain et al. (2005) provided evidence to the effect that only spontaneous episodic counterfactual thinking is impaired in patients with prefrontal damage. More recently, Van Hoeck and collaborators (Van Hoeck, Ma, Ampe, Baetens, Vandekerckhove, et al., in press; see also Van Hoeck, Ma, Van Overwalle, & Vandekerckhove 2010), asked participants to either simulate past

autobiographical events, possible future events or positive episodic counterfactual thoughts (or "upward counterfactuals", i.e., thoughts about how negative outcomes may have been better) while undergoing fMRI. Their results showed that, when compared with past and future simulations, episodic counterfactual thinking engaged prefrontal, inferior parietal, and left temporal cortices.

The present study contributes to the nascent literature on episodic counterfactual thinking by extending the findings of Addis, Pan et al. (2009) in three novel directions. First, by directly comparing brain activity associated with autobiographical memory versus neutral, positive ("upward") and negative (or "downward") episodic counterfactual thinking using spatiotemporal PLS, the present study allowed us to identify similarities and differences between regions of the core brain network associated with episodic counterfactual thinking and autobiographical recollection. Second, by manipulating outcome valence of actual episodic autobiographical events, the present study permitted us to explore similarities and differences in brain activity for positive, negative, and neutral episodic counterfactual simulations. Finally, since estimates of subjective likelihood were collected during the scans, we were able to compare brain activity associated with possible counterfactual events that were deemed as likely versus unlikely by the participants themselves.

To this end, the current study employed a variation of the experimental recombination paradigm used by Addis, Pan and collaborators in their 2009 study. During a pre-scan interview, participants recalled specific episodes in which something that they did, or that happened to them, resulted either in a positive or a negative outcome. Later, during functional magnetic resonance imaging (fMRI), participants

performed four tasks: (1) Remembered a reported autobiographical episode (*Remember* condition), (2) Imagined what would have happened if a reported event whose outcome was negative had yield instead a positive outcome (*Positive* condition), (3) Imagined what would have happened if a reported event whose outcome was positive had yield instead a negative outcome (*Negative* condition), and (4) Imagined an alternative way in which the experienced outcome could have been brought about by changing a peripheral detail of the event (*Peripheral* condition). Thus, the peripheral condition involved episodic counterfactual thinking, while keeping fixed the outcome valence of the original autobiographical event.

Three sets of data analyses were performed to investigate three distinct, yet related questions. The first question concerned whether or not episodic counterfactual thinking engages regions of the core brain network. In agreement with previous research suggesting that this network is involved in the creation of self-relevant mental simulations (Buckner & Carroll, 2007; Hassabis & Maguire, 2009; Kurczek, Duff, Jensen, Adolphs, Cohen et al., 2010), we hypothesized significant overlap between regions recruited during episodic autobiographical recollection and those recruited during episodic counterfactual thinking. Specifically, we expected to find common recruitment of core brain network regions such as the medial temporal lobe, the prefrontal cortex, and the precuneus and cingulate cortices in the parietal lobe (Spreng, Mar, & Kim, 2009).

The second question concerned whether or not there are neural differences between the experimental conditions, that is, between remembering and episodic counterfactual thinking, and also among the different forms of counterfactual simulations. As to regions preferentially recruited during recollection relative to episodic

counterfactual thinking, we expect to find greater activation in areas associated with the remembering subsystem identified by Addis, Pan et al (2009), such as posterior visual cortices, medial prefrontal cortices, and medial temporal lobe. In contrast, when it comes to regions preferentially associated with episodic counterfactual thinking, we expect to find activation in inferior parietal lobe, right temporal pole and left middle temporal gyrus, which is consistent with the results reported by Van Hoeck et al. (2010; in press). However, since—to the best of our knowledge—the present study constitutes the first attempt at manipulating outcome valence during episodic (as opposed to non-episodic) counterfactual thinking, there is little basis for advancing specific hypotheses as to which regions would be activated. At most, based upon previous results in the cognitive neuroscience of counterfactual thinking in general (Barbey et al, 2009), we expect to see differential activations in orbitofrontal cortex for positive and negative relative to peripheral counterfactuals.

Finally, the third question concerned whether or not there are neural differences between the regions engaged during episodic counterfactual thoughts rated by participants as likely versus those rated as unlikely. In their study, Addis, Pan and collaborators (2009) hypothesized that simulations requiring more imaginative work—such as imagining an unlikely alternative past event—may demand more combinatorial processing than simulations of likely alternative pasts. Consistent with this hypothesis, Weiler et al. (2010) showed that activity in right anterior hippocampus increased as a function of how subjectively unlikely simulated future events were perceived. Accordingly, we hypothesized a similar recruitment of medial temporal areas for likely relative to unlikely episodic counterfactual thoughts.

2. Materials and Methods

Participants. Seventeen healthy young adults (*M* = 21.88; SD = 3.91; range = 18 – 30; 8 female) were recruited from the University of North Carolina in Chapel Hill (UNC-CH). Participants were right-handed, native English speakers, and had no history of psychiatric or neurological disorders. Participants were recruited using flyers posted on campus, and gave written consent according to the requirements of the UNC-CH Institutional Review Board. Participants received monetary compensation. Due to a software error in the recording of ratings, data from the first two subjects were excluded from the non-rotated PLS and parametric modulation analyses. The mean-centered PLS analysis included data from all seventeen participants.

Pre-scan Stimulus Collection Interview. Approximately eight days (*M* = 8.41; SD = 1.84) prior to the scanning session, participants completed a pre-scan stimulus collection interview. They were asked to report 100 specific autobiographical memories prompted by a list of nouns adapted from Clark and Paivio (2004). They were instructed to report only event specific memories—i.e., vividly detailed recollections of single experienced events—as opposed to lifetime period or general event memories (Conway & Pleydell-Pearce, 2000); to facilitate adherence to this instruction, examples of each kind of autobiographical memory were given. One interviewer and one assistant interviewed each participant. When presented with a word (e.g., "horse"), participants reported a particular event that occurred to them at a determinate place and time (e.g., "Last summer I went horseback riding with my sister in Virginia"), and that had a specific immediate

outcome (e.g., “I fell off my horse”). All reported events were less than ten years old. Participants were asked to rate the emotional valence of each memory from 1 (“Negative Memory”) to 5 (“Positive Memory”).

Stimuli. From the reported memories collected during the interview, seventy-two were selected for stimuli. These memories were selected on the basis of two criteria. First, memories that were deemed by the two interviewers as most detailed and concrete were chosen. Second, the memories whose outcomes were most feasible to swap for alternative scenarios, in order to create episodic counterfactual events, were chosen. The alternative outcome creating the counterfactual scenario was suggested by one interviewer, and agreed upon by the second interviewer. If there was disagreement (that is, if an interviewer suggested an alternative outcome the second interviewer deemed implausible), the memory was discarded. Sixty-four of these memories were used for the experimental sessions and the remaining eight for a practice session. One fourth of these memories were positive memories (scores of 4 and 5), one fourth were negative memories (scores of 1 and 2), and the remaining half were neutral (receiving a score of 3). Three components were extracted from each memory. The *context* component was the place and time of the reported situation (e.g., “Last summer, Virginia”). The *action* component was the particular action or event the participant reported (e.g., “Horse riding”). The *outcome* component was the immediate effect occurring as a result of the reported action (e.g., “Fell off horse”). For the *control* task (see below), in which participants mentally constructed a sentence comparing sizes of different objects, we included thirty nouns referring to concrete, emotionally neutral, and imaginable objects.

Although the initial list was based on Clark and Paivio norms (Clark & Paivio, 2004), the final list included some new nouns, as extensive piloting showed that participants had very little difficulty making judgments of relative sizes among the items in the original list. Therefore, the final lists included objects for which judgments of relative sizes were not produced as rapidly. This procedure equated the means of the reaction times for the control condition with that of the experimental conditions (see Results).

Experimental tasks. In the MRI scanner, participants completed four experimental tasks: *Remember*, *Positive*, *Negative*, and *Peripheral*. All tasks had the same structure (Figure 1). Participants were presented with a screen headed by the name of the task (e.g., "Remember"), and three components: Context, action, and outcome of the event. In the *remember* task participants were presented with three components of the same neutral memory, and they were asked to remember the cued event. They were instructed to press a button as soon as they retrieved the event (i.e., at the end of the construction period). Since the components appeared on the screen for a total of ten seconds, participants were asked to elaborate on the event for any remaining time. In the *positive* condition, participants were also presented with three components, except that only the context and the action components were taken from the same reported negative memory. The outcome component was changed to be different from the reported one, in that it offered an alternative *positive* outcome to what actually happened. For example, if the participant reported a memory in which her dog underwent an unsuccessful surgery and had to be put to sleep, the outcome was changed to suggest that the surgery was successful and the dog recovered. A screen with the context, action and counterfactual positive component

appeared for ten seconds, and participants were instructed to press a button as soon as they had conjured up a counterfactual thought of this possible event. They were instructed to elaborate on that thought for the remaining time. The *negative* condition was similar to the positive condition, except that the context and action components were extracted from a positive memory, and the outcome was changed to be *negative* rather than positive. For instance, if a participant reported having won an important game of chess, the outcome was changed to suggest that she had lost. Finally, in the *peripheral* condition, the context and the outcome of the event were both extracted from the same neutral memory, but a peripheral detail of the action component was changed to suggest an alternative way in which the reported outcome could have occurred. For example, if a person reported running into her sister (outcome) while having dinner (action) at a restaurant the previous week (context), participants were instructed to imagine that instead of dinner, they were having lunch at the same restaurant.

Following the presentation of the components, participants were presented with four ratings. The first two ratings referred to the emotion and the vividness of their current mental experience. In the *emotion* rating, participants were asked to rate their current emotion from -2 (Negative) to 2 (Positive). In the *vividness* rating, participants were asked to rate how vivid the memory or their counterfactual thought appeared to them from -2 (Low vividness) to 2 (High vividness). The order of these two ratings was randomized across trials. Following a fixation cross, another set of two ratings were presented that referred to the contents of their memories or counterfactual thoughts—that is, what these mental states were about. In the *regret/relief* rating, participants were asked to rate how regretful (-2) or relieved (2) they felt for having experienced that particular

event (for the remember task), or how regretful or relieved they would have felt had the outcome turned out as suggested in the screen, as opposed to how it actually happened (for the positive, negative, and peripheral tasks). In the *likelihood* rating, participants were asked to evaluate how unlikely (-2) or likely (2) it was that the reported event had occurred. This likelihood rating was clarified for each experimental task. For the remember task, participants were asked to rate the likelihood of the event in terms of its ordinariness, with ordinary outcomes (e.g., getting food poisoning from eating expired food) being described as more likely than extraordinary events (e.g., sitting next to a celebrity at the movie theatre). For the positive and negative tasks, participants were asked to rate how likely is it that the event would have occurred with the suggested outcome as opposed to the outcome that actually happened. Finally, in the peripheral task, participants were asked to rate the likelihood that the experienced outcome would have been brought about as suggested by the alternative action as opposed to how it actually happened. Although the likelihood and regret/relief ratings for the remembering condition differ in meaning from the counterfactual conditions, such ratings were included as control ratings and will not be further analyzed here. The order to the regret/relief and likelihood ratings was randomized across trials. Each rating screen was shown for six seconds, for a total of twelve seconds per set of two trials. Jittered fixation crosses (mean duration, 4 seconds; range 2-6 seconds) were presented prior to, between, and after rating sets (see Figure 1).

Control task. To control for semantic, visual, and motor-related activity in the experimental tasks, the same control task employed by Addis and collaborators (Addis,

Pan et al., 2009) was included. In this control task, participants were presented with a slide containing three nouns. They were asked to reflect on the physical sizes of the objects referred to by those nouns, and then to mentally construct a sentence of the form "X is smaller than Y is smaller than Z", where X, Y, and Z represent the objects referred by the displayed nouns. They were asked to press a button as soon as they had covertly said the sentence to themselves. The slide with the three words was presented for ten seconds. To keep the same sequence as the experimental tasks, the same ratings used in those tasks were included after the slide containing the nouns. Participants were instructed to press a button once the rating appeared, while elaborating on the semantic definitions of the words. As such, this task controlled for similar processes as those recruited during autobiographical simulation: the retrieval and manipulation of information during construction, a task-completion decision, a motor response, and the subsequent visuo-spatial and semantic elaboration that occurs during the elaboration of mental simulations.

Scanning session. Immediately prior to scanning, the experimental and control tasks were explained to the participants, and they performed a practice session on a laptop computer to familiarize them with the tasks and the ratings. In the scanner, participants completed eight runs with eleven randomized trials per run. Each run included two trials of the remember task, two of the positive task, two of the negative task, two of the peripheral task, and three control trials. Across runs, a total of 16 remember, 16 positive, 16 negative, 16 peripheral, and 24 control trials were presented. Each run was 518 seconds long. Images were acquired in a Siemens Magnetom 3T TIM Trio MRI scanner.

Participants' heads were held in place using cushions and a headrest. An initial localizing scan was followed by high resolution sagittal and coronal T1 weighted structural scans for anatomic visualization (160 x 1 mm slices, TR=1750 ms, TE=4.38 ms), and a field map. After these structural scans, functional scans were collected during the eight experimental runs using a whole brain, gradient-echo, echo planar sequence (TR = 2 s; TE = 23; FOV = 192 mm; Matrix Size = 64 x 64; Flip Angle = 80). Interleaved slices (35 x 5 mm slices with no skip) were acquired at an angle perpendicular to the long axis of the hippocampus, as identified using the T1 structural scan. Stimuli were projected in black letters over a white background on a screen that participants could see on a mirror placed on the head coil. E-Prime software (Psychology Software Tools, Inc., Pittsburgh, PA) was used for stimulus presentation and to collect behavioral data. Responses were made on a five button MR-compatible response box.

Data analysis. Analysis of variance (ANOVA), *t*-tests and Pearson's correlations were used to analyze reaction times and ratings. Functional MRI data were preprocessed using SPM8 (Wellcome Department of Cognitive Neurology, London, UK) implemented in MATLAB (Mathworks, Natick, MA). Images were re-oriented, slice-time corrected, realigned and unwarped, normalized to the Montreal Neurological Institute (MNI) template (resampled at 2 x 2 x 2mm voxels) and spatially smoothed using a 8mm full-width half maximum isotropic Gaussian kernel.

Next, two spatiotemporal PLS analyses were performed. First, a hypothesis-free mean-centered analysis was conducted with a 20-second temporal window (10 TRs). Previous similar studies have employed different temporal windows depending on the

length of the trial, ranging from 12-seconds (Addis, Pan et al., 2009) to 20-seconds (Addis, Roberts, & Schacter, 2011; Addis, Knapp, Roberts, & Schacter, 2012). Given the length of the trials in the current study (42-seconds), a 20 second (10 TR) window was chosen. Each subject's fMRI data were entered into a data matrix (or datamat) with onset times specified as 2000 ms after stimulus onset (to allow for reading of the stimuli). Participants' datamats were then cross-correlated with a matrix of vectors coding for the tasks (design matrix). Singular value decomposition is then applied to reveal latent variables (LVs) that can best account for the most covariance. Since the number of LVs is constrained by the degrees of freedom, four LVs explaining the most covariance were identified. The statistical significance of each LV was established by computing 500 permutation tests on the singular values to determine the probability that the permutated singular value exceeds the singular value of the original latent variable. LVs for which the probability is $p < .05$ are thus considered significant. To identify voxels that reliably contributed to each LV, a bootstrap estimation method was computed 300 times. Clusters larger than 100 mm^3 comprising voxels with a ratio of the salience to the bootstrap standard error values (i.e., the "bootstrap ratio"; BSR) greater than 3.2 ($p < .0003$) are reported. PLS identifies whole brain patterns of activity in a single analytic step, thus, no correction for multiple comparisons is required. The local maximum for each cluster was defined as the voxel with a BSR higher than any other voxel in a 2cm cube centered on that voxel.

Second, a non-rotated PLS analysis was performed in order to examine the differences in brain activity between likely and unlikely episodic counterfactual thinking.

Since we had the a priori hypothesis that the pattern of brain activity for likely counterfactuals would be more similar to that of remember than unlikely counterfactuals, a non-rotated analysis allowed us to enter specific contrasts of interest: first, a contrast of remember and likely relative to unlikely and control; and second, a contrast of remember, likely and unlikely relative to the control task. To that end, we split the trials of both the Positive and the Negative conditions into two groups, *Likely* and *Unlikely*, according to the participant's likelihood rating. In the likely condition, we included positive and negative trials to which participants gave a rating of 4 or 5, and in the unlikely condition we included those to which participants gave a rating of 1 or 2. All other trials were excluded. We then created a datamat with four conditions: remember, likely, unlikely and control. As in the mean-centered analysis above, 500 permutations were conducted and for each LV a bootstrap estimation was carried out 300 times (BSR = 3.2, $p < .0003$).

Finally, since the non-rotated analysis explained above bins together trials from different experimental conditions (i.e., positive and negative counterfactuals) into post-hoc likely and unlikely conditions, it would still be unclear whether brain activity associated with each experimental condition is differentially influenced by ratings of likelihood. Consequently, and based upon previous research showing differential increases and decreases in brain activity associated with perceived likelihood during episodic future thinking (Weiler et al., 2010), we decided to carry out both negative and positive parametric modulations for each experimental condition of the current study. All parametric modulations were conducted in SPM8. At the fixed-effects level, condition regressors for positive, negative and peripheral counterfactual conditions were specified. For each, the likelihood rating was entered as a parametric modulation covariate. Because

vividness ratings showed a significant correlation with likelihood ratings, the ratings for vividness were entered as an additional (nuisance) parametric modulation covariate in order to isolate the orthogonal contribution of likelihood for the modulation of brain activity. All covariates were modeled linearly. For each condition, contrasts were computed to identify regions in which activity was positively or negatively correlated with likelihood; this was done separately for each condition. Finally, at the random-effects level, each set of contrast images (e.g., contrast images from all subjects that coded for a positive parametric modulation in a particular condition) were entered into a one-sample *t*-test to determine which regions correlated with likelihood at the group level. Following previous parametric modulation studies (Rombouts, Scheltens, Machielsen, Barkhof, Hoogeraad et al, 1999; Addis, Moscovitch, Crawley, & McAndrews, 2004; Addis & Schacter, 2008) a threshold was set for $p < .005$, $k = 10$.

3. Results

3.1. Behavioral results.

Behavioral results. Only trials for which participants gave a rating are included in all analyses. On average, each participant contributed 14.4 (SD=2.13) memories for the remember condition, 14.8 (SD=2.17) positive episodic counterfactuals, 14.67 (SD=2.18) negative counterfactuals and 14.36 (SD=2.34) peripheral counterfactuals.

RTs did not differ across conditions ($F(4,80) = .816$, $p = .52$). The mean RTs across conditions were as follows: remember M = 4063ms (SD = 956 ms), positive M = 4331ms (SD = 1170 ms), negative M = 4326ms (SD = 1074 ms), peripheral M = 4512ms (SD = 1177 ms), and control M = 4781ms (SD = 1586 ms).

On a 1-to-5 scale for *emotion*, positive episodic counterfactuals were rated as more positive (M = 4.13; SD = .44) than negative (M = 2.88; SD = .044, $t(28) = 7.82$, $p <$ .005; alpha level set at p<.05) and peripheral counterfactuals (M = 3.45; SD = .50, $t(28) =$ 3.96, $p < .005$), but they were not significantly different from the remember condition (M = 3.91; SD = .45, $t(28) = 1.35$, $p = .18$). For *vividness,* the ratings for remembering (M = 3.90; SD = .36) were significantly higher than both peripheral (M = 3.46; SD = .48, $t(28)$ = 2.88, $p < .05$) and negative counterfactuals (M = 2.65; SD = .35, $t(28) = 9.76$, p < .005). Likewise, positive counterfactuals were rated as more vivid (M = 4.02; SD = .98) than both peripheral ($t(28) = 3.16$, p < .005) and negative counterfactuals ($t(28) = 8.80$, p < .005). Participants rated the *likelihood* of negative counterfactuals (M = 2.31; SD = .32) as significantly lower than positive (M = 3.62, SD = .50; $t(28) = 8.50$, p < .005), and peripheral counterfactuals (M = 3.24, SD = .52; $t(28) = 5.88$, p < .005). Finally, for the *relief/regret* ratings, participants tended to say that they would have felt more relieved had the outcomes of negative events been positive (M = 3.78; SD = .52) than they would, had the outcomes of positive events being negative (M = 2.40; SD = .40; $t(28) = 8.06$, p < .005). The *relief/regret* ratings for peripheral counterfactuals (M = 3.31; SD = .54) were also significantly different from positive ($t(28) = 2.53$; p = .02) and negative counterfactuals ($t(28) = 6.00$; p < .005). Additionally, across counterfactual conditions, there were strong and significant correlations between the ratings for likelihood and vividness (r = .53, p < .001) and also between the ratings for emotion and regret/relief (r = .50, p < .001). Finally, there were weak but significant correlations between the ratings for emotion and vividness (r = .11, p < .005), and also between the ratings for vividness and regret (r = .18, p < .005). No other effects were significant.

Behavioral results related to the non-rotated PLS analysis. As mentioned above, for the non-rotated analysis we split the trials from the positive and the negative conditions into two post-hoc conditions, likely and unlikely counterfactuals, according to likelihood ratings. Trials that received ratings of 4 and 5 only were included in the likely condition, whereas trials that received ratings of 1 and 2 only were included in the unlikely condition. All other trials were excluded. On average, each participant contributed 14.4 (SD = 2.13) episodes for the remembering condition, 11.20 (SD = 3.61) for the likely condition, and 12.53 (SD = 3.66) for the unlikely condition. The average rating of likelihood for the likely condition was 4.46 (SD = .50) while for the unlikely condition was 1.58 (SD = .49). A *t*-test revealed that these means were significantly different ($t(167) = 115.7$; $p < .001$). After the split, the averages for the remaining ratings for both groups were as follows. Likely counterfactuals were rated as more positive (M = 3.68; SD = 1.27), more vivid (M = 4.10; SD = .95) and more relieving (M = 3.28; SD = 1.40) than unlikely counterfactuals (emotion: M = 3.49; SD = 1.37; vividness: 2.61; SD = 1.36; regret/relief: M = 2.83; 1.42). In addition to the ratings of likelihood, likely and unlikely counterfactuals also differed in their ratings of vividness ($t(167) = 48.1$; $p < .001$). This result was expected given the correlation reported above. No other effects were significant.

3.2. Mean-centered PLS analysis. This analysis identified one significant latent variable (LV1; $p < .03$; singular value = 133.53), which explained 52.51% of the crossblock covariance. This LV differentiated the remember, positive and negative conditions from the control condition, with the peripheral condition not contributing to the pattern

(indicated by the error bars for this condition crossing zero; figure 2A). The regions with negative saliences (Table 1) were commonly engaged by the remember, positive and negative counterfactuals tasks relative to the control task. This neural differentiation was evident as early as the second TR (Figure 2B). The set of activated regions (see Tables 1 and 2, and Figures 2C and 2D) included the bilateral cingulate cortex (BA 23 and 24) and superior temporal gyri (BA 38), as well as right inferior frontal gyrus (BA 9), right parahippocampal gyrus (BA 28) extending into the hippocampus, right precentral gyrus (BA 6), left superior parietal/precuneus (BA 7) and left middle temporal lobe (BA 39 and 21). As hypothesized, these regions are all part of the core network identified in previous studies of remembering and imagining (Schacter, Addis, & Buckner, 2008).

In addition, and similar to Addis, Pan et al's (2009) reported results, the laterality of some identified regions varied across the duration of the trials. Right regions showed activity only during the first 4 TRs, after which, all activity occurred in the left hemisphere. However, the laterality of some regions never varied. For instance, the activity of the posterior cingulate cortex (BA 23) appeared in the left hemisphere, whereas the parahippocampal gyrus (BA 28) and the inferior frontal gyrus (BA 9) showed only right activations. The laterality of yet other regions shifted over the course of the trial. For example, the cerebellum showed left activity during TR 2 but during TR 4 the activity was in the right hemisphere. Similarly, the precuneus (BA 7) showed right activity during TR 3 but it shifted to the left during TR 4. Finally, while the superior temporal gyrus (BA 38) was active during most of the trial, its activity was preferentially in the left hemisphere, except during TR 4 were it was bilateral.

Regions with positive saliences were associated with the control condition. These regions included bilateral middle frontal gyrus (BA 46 and 10), superior frontal gyrus (BA 9 and 8) and claustrum, as well as right medial frontal gyrus (BA 25), right supramarginal gyrus (BA 40) and left superior temporal gyrus (BA 38). For brevity, regions associated with the control task are not reported in Table 1.

3.3. Non-Rotated PLS analysis. We conducted a non-rotated analysis to explore differences in patterns of brain activity between counterfactuals that were deemed as likely or unlikely by the participants, and the retrieved memories from the remembering condition (trials from the counterfactual conditions were binned into these post-hoc conditions, as explained above in Section 3.1). This analysis revealed a significant LV (LV 2; $p < .02$; Singular Value 167.95), which explained 38.26% of the crossblock covariance. This LV differentiated activations for remember and likely versus unlikely and control (positive versus negative salience's respectively; see tables 2 and 3, and Figure 3). LV 2 indicates significant overlap in the brain regions engaged during remembering and likely episodic counterfactual thinking, including bilateral inferior parietal lobe (BA 40), right precuneus/superior parietal lobule (BA 7) and superior temporal gyrus (BA 38), as well as left superior frontal gyrus (BA 9 and 10), middle and inferior frontal gyrus (BA 45, 46 and 47), middle and inferior temporal gyri (BA 21 and 22), angular gyrus (BA 39) and parahippocampal gyrus (BA 36). An extensive set of regions were active during unlikely episodic counterfactual thoughts and the control condition (Table 3), including bilateral superior frontal gyrus (BA 6, 8 and 9), middle frontal gyrus (BA 10 and 11), cingulate (BA 24, 29, 31 and 32), middle temporal gyrus

(BA 22), parahippocampal gyrus (BA 19 and 30), putamen and precuneus. Additionally, it included left inferior parietal lobule (BA 40), caudate and insula (BA 13), as well as right superior temporal gyrus (BA 42 and 38), parahippocampal gyrus (BA 27), thalamus and hypothalamus. Of note, the laterality of some of the regions shifted across the duration of the trial. For instance, the parahippocampal gyrus (BA 19) showed activations at TR 3 (4 to 6 seconds after stimulus onset) and then it showed contralateral activation at TR 4 (6 to 8 seconds after stimulus onset). Likewise, the middle frontal gyrus (BA 10) showed left activity in TR 3, then it showed right activity at TR 5, and then it showed bilateral activity at TR 6 (10 to 12 seconds after stimulus onset).

Of note, the second contrast in this non-rotated analysis showed a strong trend toward differentiating remember, likely and unlikely counterfactuals from the control condition ($p = .06$; Singular Value 167.84). Presumably, this LV did not reach significance because the likely and unlikely conditions included trials from the peripheral condition which, as indicated by LV 1, did not contribute to the overall activation of regions from the core network. However, to highlight the similarities between this LV (LV 3) and LV 1, the regions of activation revealed by this LV are included in supplementary materials. Finally, to verify that the difference in brain activity patterns revealed by LV 2 for likely and unlikely counterfactuals was not driven by the control condition, a non-rotated analysis including only remember, likely, and unlikely conditions was conducted. This analysis revealed a significant LV (LV 4; $p < .05$, Singular Value 143.0), which explained 39.04% of the crossblock covariance. Although some new regions emerged relative to LV2, overall the spatial patterns were overlapping. This result confirms that the differential pattern of activation revealed by LV 2 was

driven by a difference in perceived likelihood of counterfactuals, rather than by the control condition. The regions of activation revealed by LV 4 are included in the Supplementary Materials.

3.4. Parametric modulation. Results of the parametric modulation analyses are reported in Table 4. Regions whose activity increased linearly as the ratings for likelihood increased included left inferior and middle frontal gyrus for positive counterfactuals, anterior left parahippocampal gyrus and right hippocampus for negative counterfactuals, and superior temporal, cingulate and posterior left parahippocampal gyri for peripheral counterfactuals. Regions whose activity decreased linearly as the rating for likelihood increased included thalamus, superior temporal and right parahippocampal gyrus for positive counterfactuals, middle temporal and angular gyrus for negative counterfactuals, and post-central and right superior frontal gyrus for peripheral counterfactuals.

4. Discussion

4.1. Behavioral results. Our results indicate that, overall, participants tended to conjure up more vivid positive episodic counterfactual thoughts than negative counterfactual thoughts. In other words, thoughts about how things could have turned out for the better were experienced by participants as more vivid and detailed than thoughts about how things could have turned out for the worse. This result is consistent with previous evidence indicating that people tend to remember their past, and also simulate their future, with a positivity bias (Bower, 1981; Holmes, 1970; Meltzer, 1930; Sharot, Riccardi, Raio, & Phelps, 2007; Szpunar, Addis, & Schacter, 2012; Waldfogel, 1948;

Walker & Skowronski, 2009; Wilbur, Skowronski, Thomson, 2003). It appears, therefore, that this positivity bias carries over to episodic counterfactual thinking. Additionally, participants thought that possible positive outcomes to experienced negative events were more likely than possible negative outcomes to experienced positive events. In other words, participants indicated that it was more likely that a negative event could have had a positive outcome than a positive event could have had a negative outcome. This result is also consistent with other well-established evidence on "optimism bias", according to which people tend to overestimate the likelihood of a possible event happening over a negative event (Hoch, 1984; Sharot et al., 2007; Weinstein, 1980). The fact that episodic counterfactual thinking is subject to biases known to affect episodic recollection and future thinking strongly suggests the involvement of shared cognitive processes. Finally, the strong correlation between the ratings of vividness and likelihood indicates that more likely episodic counterfactual thoughts are experienced more vividly and with more detail. From the point of view of the constructive episodic simulation hypothesis, this result suggests that more likely counterfactuals conjure up more episodic details than unlikely counterfactuals, which would explain why we experience the former more vividly than the latter.

4.2. Episodic counterfactual thinking and the core brain network. The first question explored in the present study concerned whether or not episodic counterfactual thinking, independent of the subjective experience of likelihood, engages regions of the core brain network. Previous research has shown that cognitive tasks involving self-projection—predominantly autobiographical recollection and episodic future thinking—tend to recruit

a common network of brain regions. Since episodic counterfactual thinking also involves self-projection, we hypothesized that this task would also recruit these brain regions as well. The results of the mean-centered PLS analysis lend strong credence to this hypothesis. The identification of a latent variable (LV 1) differentiating the remembering and the episodic counterfactual tasks from the control task suggests that there is significant neural overlap between brain regions engaged during autobiographical recollection and those engaged during episodic counterfactual thinking (Figure 2; see also LV 3 in the Supplementary Materials for similar findings). Moreover, essentially all the regions associated with the experimental tasks identified by LV 1 are part of the core brain network. Of special interest is the right medial temporal activation in TR 2; although this cluster peaked in the parahippocampal gyrus, it extended to anterior regions of the hippocampus proper (y = -22). Activation of right hippocampus and parahippocampal gyrus during the construction of episodic simulations has been a consistent result in the episodic memory and future thinking literature (Okuda et al., 2003; Szpunar et al., 2007, Hassabis et al., 2007; Weiler et al., 2010; Addis et al., 2007, 2009, 2011). Our results are thus in agreement with these findings, showing that the right hippocampus and adjacent regions in the parahippocampal gyrus are recruited during the simulation of episodic counterfactual thoughts as well. This pattern of activity is also consistent with the constructive episodic simulation hypothesis, and lends credence the claim that right hippocampus may be supporting the relational processing (Schacter & Wagner, 1999; Eichembaum, 2001; Giovanello, Schnyer, & Verfaellie, 2004; Chua, Rand-Giovennetti, Schacter, Albert, & Sperling, 2007) required to flexibly recombine previously encoded memories into novel episodic counterfactual thoughts. More recent

data indicate a role for right hippocampus in encoding future simulations for later recall (Martin, Schacter, Corballis, & Addis, 2011), and it is possible that the present findings also reflect, to some extent, activity related to the encoding of a novel counterfactual simulation into episodic memory.

The significant neural overlap between the remembering and the episodic counterfactual tasks also supports the claim that the mechanisms underlying episodic autobiographical recollection contribute not just to future simulation but to other forms of imagination including our capacity to think of alternative ways in which our past could have occurred. Recent research in the cognitive neuroscience of decision-making provides further evidence in favor of this claim. In studies designed to elicit participants' feelings of regret associated with counterfactual thoughts regarding what would have happened had they made a different choice (e.g., *regret gambling task;* Camille, Coricelli, Sallet, Pradat-Diehl, Duhamel et al, 2004), a consistent result is the involvement of critical regions within the core brain network—specifically, the middle frontal gyrus, the orbitofrontal cortex, the anterior cingulate cortex, and the hippocampal and parahippocampal regions (e.g., Coricelli, Critchley, Joffily, O'Doherty, Sirigu et al., 2005; Coricelli, Dolan, & Sirigu, 2007). Additionally, our results are consistent with the aforementioned findings by Van Hoeck et al's (in press; also Van Hoeck et al., 2010) in which regions of the core brain network were active while participant's freely engaged in upward episodic counterfactual thinking. Taken together, these results, as well as the results from our non-rotated PLS analysis (discussed below), strongly suggest that similar neural mechanisms underlie episodic autobiographical memory and episodic counterfactual thinking – particular when the counterfactual thoughts are likely and

plausible. Our capacity to imagine alternative ways in which our past could have occurred appears to involve similar processes to those employed when we think about what in fact occurred in our pasts.

Finally, it should be noted that, although the current study was also designed to shed light on whether or not there were differences in brain regions between different valenced counterfactual simulations, the mean-centered analysis did not yield a significant LV differentiating patterns of activations between such conditions. This result could be due to many factors. One possibility is simply lack of power, since the number of mental simulations that were emotionally neutral (in the remembering and peripheral conditions) were double the number of those that had either a negative or a positive valence. Another possibility is that the suggested outcomes used in the manipulation were not strong enough to elicit significant emotional effects. As such, future research may benefit from using more emotionally loaded stimuli in during counterfactual generation, as well as a larger number of observations, in order to clarify the differential contribution of the core brain regions in episodic counterfactual thinking.

4.3. Likely versus unlikely episodic counterfactuals thoughts. The current study was also designed to investigate neural differences between the regions engaged during likely versus unlikely episodic counterfactual thoughts. Even though participants were only presented with counterfactual outcomes considered plausible by both interviewers, it is still possible that some of those outcomes are not considered plausible or likely by the participants themselves. Indeed, the fact that roughly a third of the counterfactuals suggested to the participants were judged as unlikely confirms this claim. The non-

rotated PLS analysis conducted on the sets of likely and unlikely counterfactual thoughts yielded a significant variable, LV 2, which differentiated between remembering and likely counterfactuals, on the one hand, and unlikely counterfactuals and control, on the other (Figure 3). LV 2 showed substantial overlap in the pattern of regions activated during remembering and likely episodic counterfactual thoughts, suggesting greater involvement of the remembering subsystem in our capacity to think of possible past events that we think are more likely to have happened in our past relative to possible past events we think are unlikely that could have occurred.

As mentioned before, Addis, Pan and collaborators (2009) suggested that unlikely episodic counterfactual thoughts may require more "imaginative work" than likely ones, which in turn would be reflected in greater activation of regions associated with the imagining subsystem. Our results partially confirm this prediction. As shown in figure 4 (included in supplementary materials), simulating unlikely counterfactuals engaged some regions of the imagining subsystem, such as the parahippocampal cortex and the inferior frontal gyrus, to a larger extent than does both remembering and simulating likely counterfactuals. Similarly, left superior frontal gyrus (BA 10) is sensitive to the likelihood of counterfactuals, showing a significant decrease in signal during the construction of unlikely episodic counterfactual thoughts while remaining at baseline during the construction phases in remembering and likely counterfactual thinking. However, given the substantial overlap between unlikely counterfactuals and the control condition in LV 2—which, as indicated by the temporal brain scores (Figure 3B), remained tightly coupled for at least the first 6 TRs—it is unclear whether these processes reflect an increase in combinatorial process of episodic details or some other imagination-

related capacity also deployed during the control task. Further research contrasting episodic counterfactual thinking with purely imaginative tasks could help disentangle this issue.

The results of the parametric modulation analysis also shed light on the involvement of certain regions of the core brain network—particularly frontal regions, the cingulate gyrus, and the medial temporal lobe—during the construction of likely and unlikely counterfactual simulations. Increased likelihood during the construction of positive or "upward" episodic counterfactual thoughts—that is, thoughts in which participants were instructed to change the negative outcome of an actual personal experience for an alternative positive one—correlated with increased activity in left middle and inferior frontal gyrus (BA 10 and 46). Similarly, the increased likelihood of negative or "downward" episodic counterfactual thoughts, whereby participants thought of alternative negative outcomes to memories of events with actual positive outcomes, was also associated with increased activity in middle and inferior frontal gyrus (BA 9, 46 and 47), but in the right hemisphere. Previous research has shown that some of these regions—particularly, BA 10 and 46—are activated during self-referential simulations in both episodic memory and episodic future thinking (Okuda et al., 2003; Addis et al., 2007; D'Argembeau, Raffard, & Van der Linden, 2008). These results are also consistent with research on counterfactual thinking showing that medial prefrontal cortex is preferentially active when participants think about alternative scenarios involving themselves versus others (Belderrain et al., 2005), while inferior regions of the orbitofrontal cortex—particulary BA 9, 10 and 47—have been found to be involved during upward and downward counterfactual thinking (Elliot, Dolan, & Frith, 2000;

Coricelli et al., 2005; Barbey et al., 2009). The results of the present parametric modulation are consistent with these observations, insofar as they show that the subjective likelihood of both positive and negative episodic counterfactual thoughts—which by definition are self-referential—modulate regions in the prefrontal cortex expected to be engaged by upward and downward counterfactual thoughts about oneself. This observation suggests that as the perceived likelihood of our episodic counterfactual thoughts increases, the self-referential quality of these simulations might also increase.

In addition to frontal regions, increases in ratings of likelihood for negative counterfactuals were also associated with increased activity in anterior cingulate cortex, which has been consistently identified as part of the core brain network (Buckner & Carroll, 2007; see below). This result is consistent with previous studies showing the involvement of the anterior cingulate cortex in general counterfactual thinking (Coricelli et al, 2005), as well as episodic simulations irrespective of temporal dimension (Addis et al., 2009; D'Argembeau, Xue, Lu, Van del Linden, & Bechara, 2008). Similarly, the insula, which has reciprocal connections with the medial prefrontal and anterior cingulate cortices, has been associated with the regulation of evaluative and affective processes (Bertson, Norman, Bechara, Bruss, Tranel et al., 2011). In particular, increases in insula activation have been correlated with increases in emotions associated with risk-averse behavior, particularly regret (e.g., Xue, Lu, Levin, & Bechara, 2010). However, a recent study investigating common and unique neural activations for autobiographical, episodic, and semantic retrieval, showed insula activation to be preferentially associated with autobiographical recollection (Burinova & Grady, 2007). As such, the positive modulation of insula observed in our studies may reflect either the evaluative processes

in the emotion associated with likely episodic counterfactual thinking—which would suggest that as the subjective likelihood of our counterfactual thoughts increases so does our evaluative affective processing—or the contribution of autobiographical recollection to the construction of mental simulations about alternative ways our personal past could have been. Further research may be needed to clarify why it is that regions like insula and the anterior cingulate cortex are differentially engaged depending on the emotional direction of the counterfactual simulation. Likewise, future research may be able to shed light on the fact that subjective likelihood increased activity in more brain regions for negative and peripheral counterfactual simulations whereas more brain regions showed decreases in activity during positive counterfactual simulations.

Of special interest is the incremental activation of regions in the temporal lobe as a function of increased subjective likelihood in negative and peripheral episodic counterfactual thinking. For both negative and peripheral counterfactuals, likelihood positively modulated activity in bilateral superior temporal (BA 22 and 38) and left parahippocampal gyri (BA 19 and 30), regions also revealed to be associated with remembering and likely counterfactual thoughts in LV 2. In their study of occurrence probability for episodic future thinking, Weiler and colleagues (2010) report a positive modulation of BA 22 for richness of details. However, in the present study, we controlled for vividness and still found BA 22 to be positively modulated by increased subjective probability. As such, activity in BA 22 appears to increase during the construction of imagined scenarios as a function of their perceived subjective likelihood independently of the richness of the details with which such scenarios are experienced. Additionally, increased subjective probability also modulated activation in the left parahippocampal

gyrus. Activations in this area are common during construction periods in episodic future thinking tasks. For example, Szpunar and collaborators (2009) found greater activity in parahippocampal gyrus during the construction of possible future events in familiar versus unfamiliar settings. Future research could investigate whether the incremental recruitment of the parahippocampal cortices in episodic counterfactual thoughts perceived as more vivid and likely to have happened is related to the retrieval of familiar contexts.

As noted above, the increment in activity in right hippocampus as a function of increased subjective likelihood in negative episodic counterfactual thoughts is also noteworthy. This result appears inconsistent with Weiler et al's (2010) parametric modulation of occurrence probability in episodic future thinking. They reported a *decrease* in activity in right hippocampus, correlated with increase subjective likelihood, in a cluster peaking within ~4 mm of the cluster's peak reported here. In their paper, Weiler and colleagues interpret their result as suggesting that the decrease in right hippocampal activity—strongly associated with binding of novel associations during episodic simulation (see Addis & Schacter, 2012, and Schacter & Addis, 2009, for a review)— may reflect a higher demand during the binding of imagined future probable events relative to improbable events; since the former tend to be more similar to past events, they may require less episodic recombination as their components are usually experienced together. However, there are reasons to believe that neither the result reported here, nor our interpretation of it, is actually inconsistent with Weiler et al's result. For one, they explored hippocampal activity during event elaboration, not during event construction, as the current study does. It may be that the right hippocampus is

differentially responsive to likelihood during these difference phases of simulation. Second, their study did not differentiate between emotionally positive and negative episodic future thoughts, while here we link right hippocampal activity specifically to negative counterfactual thoughts. Although further research is needed to pin down the differential contributions of the hippocampus to both downward and upward counterfactual thinking, it is possible that the modulation of hippocampal activity by likely counterfactuals we found in the current study is actually driven by the kind of emotion associated to the simulation, rather than the mere experience of subjective probability.

5. Conclusion.

Consistent with previous studies, our results suggest that episodic counterfactual thinking engages many of the same brain regions that form the core network recruited during episodic autobiographical recollection. In addition, several of the recruited areas were differentially modulated by the participant's own perception of how likely it is that a certain event could have occurred in a particular way. Moreover, our analysis also revealed that there is more overlap between the neural regions associated with episodic recollection and likely episodic counterfactual thoughts than with counterfactual thoughts considered unlikely. Taken together, these results suggest that similar cognitive processes may support episodic counterfactual thinking and episodic autobiographical recollection, particularly when counterfactual thinking is realistic. In accordance with the constructive episodic simulation hypothesis, our results partially support the claim that activity in the hippocampus and adjacent medial temporal areas, as well as medial and lateral frontal

regions, may reflect the increased involvement of recombinatory processes in the construction of unlikely episodic counterfactual thoughts. However, the large overlap in activation between unlikely counterfactuals and the control condition makes it difficult to draw strong conclusions about the precise process associated with such activations. Nonetheless, taken together, our results are consistent with recent work on episodic future thinking, indicating that episodic memory not only provides the informational components of our thoughts about what happened in the past but also about what may happen in the future and what might have happened in the past. Exploring the constraints imposed by episodic memory to our thoughts about what may happen in the future and what may have happened in the past offers a rich avenue for future research.

Acknowledgments

This work was supported by National Institute on Aging grant AG028774 awarded to K.S.G., National Institute of Mental Health grant MH060941 awarded to D.L.S., and Royal Society of New Zealand Marsden Grant (UOA0810) and Rutherford Discovery Fellowship (RDF-10-UOA-024) awarded to D.R.A.

References

Addis, D. R., Moscovitch, M., Crawley, A. P., McAndrews, M. P. (2004). Recollective qualities modulate hippocampal activation during autobiographical memory retrieval. *Hippocampus* 14, 752-762.

Addis, D.R., Wong, A.T., & Schacter, D.L. (2007). Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. *Neuropsychologia* 45, 1363-1377.

Addis, D.R., Wong, A.T., & Schacter, D.L. (2008). Age-related changes in episodic simulation of future events. *Psychological Science* 19: 33-41.

Addis, D.R., & Schacter, D.L. (2008). Effects of detail and temporal distance of past and future events on the engagements of a common neural network. *Hippocampus.* 18: 227-237.

Addis, D.R., Pan, L., Vu, M., Laiser, N., & Schacter, D.L. (2009). Constructive episodic simulation of the future and the past: Distinct subsystems of a core brain network mediate imagining and remembering. *Neuropsychologia*. 47: 2222-38.

Addis, D. R., Sacchetti, D. C., Ally, B. A., Budson, A. E., & Schacter, D. L. (2009). Episodic simulation of future events is impaired in mild Alzheimer's disease. *Neuropsychologia, 47,* 2660-2671.

Addis, D. R., Musicaro, R., Pan, L., & Schacter, D. L. (2010). Episodic simulation of past and future events in older adults: Evidence from an experimental recombination task. *Psychology and Aging*. 25, 369-376.

Addis, D. R., Cheng, T., Roberts, R., & Schacter, D.L. (2011). Hippocampal contributions to the episodic simulation of specific and general future events. *Hippocampus*. 21, 1045-1052 .

Addis, D.R., Roberts, R.P. & Schacter, D.L. (2011). Age-related neural changes in autobiographical remembering and imagining. *Neuropsychologia*. 49: 3656-3669.

Addis, D.R., Knapp, K., Roberts, R.P., & Schacter, D.L. (2012). Routes to the past: Neural substrates of direct and generative autobiographical memory retrieval. *NeuroImage*. 59, 2908-2922.

Addis, D.R. & Schacter, D.L. (2012). The hippocampus and imagining the future: Where do we stand? *Frontiers in Human Neuroscience*. *5*, 173. doi: 10.3389/fnhum. 2011.00173

Atance, C. M., & O'Neill, D. K. (2001). Episodic future thinking. *Trends in Cognitive Science*. 5, 533-539.

Beldarrain M., Garcia-Monco J., Astigarraga E., Gonzalez A., & Grafman J. (2005). Only spontaneous counterfactual thinking is impaired in patients with prefrontal cortex lesions. *Cognition and Brain Research*. 24, 723–726

Barbey, A.K., Krueger, F., & Grafman, J. (2009). Structured event complexes in the medial prefrontal cortex support counterfactual representations for future planning. *Philosophical Transactions of the Royal Society of London. B: Biological Sciences*. 364, 1291-1300.

Berntson G.G, Norman G.J., Bechara, A., Bruss, J., Tranel D., & Cacioppo J.T. (2011). The insula and evaluative processes. *Psychological Science*. 22, 80-6.

Bower, G. H. (1981). Mood and memory. *American Psychologist.* 36, 129-148.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Science*. 11, 49–57.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Science.* 1124, 1–38.

Burianova, H., & Grady, C.L. (2007). Common and unique neural activations in autobiographical, episodic, and semantic retrieval. *Journal of Cognitive Neuroscience.* 19, 1520-1534.

Byrne, R.M.J. (2002). Mental models and counterfactual thoughts about what might have been. *Trends in Cognitive Science*. 6, 426-431.

Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J., & Sirigu, A. (2004). The Involvement of the Orbitofrontal Cortex in the Experience of Regret. *Science*. 304, 1167-1170.

Chua, E., Rand-Giovennetti, E., Schacter, D.L., Albert, M.S., & Sperling, R.A. (2004). Dissociating confidence and accuracy: Functional magnetic resonance imaging shows origins of subjective memory experience. *Journal of Cognitive Neuroscience.* 16, 1131–1142.

Clark, J.M., & Paivio, A. (2004). Extensions of the Paivio, Yuille, and Madigan (1968) norms. *Behavioral and Research Methods.* 36, 371–383

Conway, M.A., & Pleydell-Pearce, C.W. (2000). The search of autobiographical memories in the self-memory system. *Psychological Review.* 107, 261-288.

Coricelli, G., Critchley, H.D., Joffily, M., O'Doherty, J.P., Sirigu, A. & Dolan, R.J., (2005). Regret and its avoidance: a neuroimaging study of choice behavior. *Nature Neuroscience*. 8, 1255-1262.

Coricelli, G., Dolan, R.J, & Sirigu, A. (2007). Brain, emotion, and decision making: the paradigmatic example of regret. *Trends in Cognitive Science.* 11, 258-265.

Crovitz, H.F., & Schiffman, H. (1974). Frequency of episodic memories as a function of their age. *Bulletin of the Psychonomics Society*. 4, 517-518.

D'Argembeau, A., & Van der Linden, M. (2004). Phenomenal characteristics associated with projecting oneself back into the past and forward into the future: Influence of valence and temporal distance. *Consciousness and Cognition*. 13, 844–858.

D'Argembeau, A., & Van der Linden, M. (2006). Individual differences in the phenomenology of mental time travel: The effects of vivid visual imagery and emotion regulation strategies. *Consciousness and Cognition*. 15, 342–350.

D'Argembeau, A., Xue, G., Lu, Z. L., Van der Linden, M. & Bechara, A. (2008). Neural correlates of envisioning emotional events in the near and far future. *Neuroimage*. 40, 398–407.

D'Argembeau, A., Raffard, S., & Van der Linden, M. (2008). Remembering the past and imagining the future in schizophrenia. *Journal of Abnormal Psychology.* 117, 247–251.

D'Argembeau, A.D., Stawarczyk, D., Majerus, S., Collette, F., Van der Linden, M., Feyers, D., Maquet, P., & Salmon, E. (2009). The neural basis of personal goal processing when envisioning future events. *Journal of Cognitive Neuroscience.* 22, 1701–1713.

De Brigard, F., & Giovanello, K. (2012). Influence of outcome valence in the subjective experience of episodic past, future, and counterfactual thinking. *Consciousness and Cognition*. 21, 1085-1096.

De Brigard, F., Szpunar, K.K., & Schacter, D.L. (in press). Coming to grips with the past: Effects of repeated simulation on the perceived plausibility of episodic counterfactual thoughts. *Psychological Science*.

Dickson, J. M., & Bates, G. W. (2005). Influence of repression on autobiographical memories and expectations of the future. *Australian Journal of Psychology*. 57, 20–27.

Eichenbaum, H. (2001). The hippocampus and declarative memory: Cognitive mechanisms and neural codes. *Behavioral Brain Research*. 127, 199–207.

Elliott R., Dolan R.J., & Frith C.D. (2000). Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cerebral Cortex*. 10, 308–317.

Epstude, K., & Roese, N.J. (2008). The Functional Theory of Counterfactual Thinking Personal. *Social Psychological Review.* 12: 168-192.

Gaesser, B., Sacchetti, D. C., Addis, D. R., & Schacter, D. L. (2011). Characterizing age-related changes in remembering the past and imagining the future. *Psychology and Aging*. 26, 80-84.

Gamboz, N., De Vito, S., Brandimonte, M.A., Pappalardo, S., Galone, F., Iavarone, A., & Della Sala, S. (2010). Episodic future thinking in amnesic mild cognitive impairment. *Neuropsychologia*. 48, 2091-2097.

Giovanello, K.S., Schnyer, D., & Verfaellie, M. (2004). A critical role for the anterior hippocampus in relational memory: Evidence from an fMRI study comparing associative and item recognition. *Hippocampus*. 14, 5-8.

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Science U. S. A*. 104, 1726–1731.

Hassabis D, Kumaran D., & Maguire E.A. (2007). "Using imagination to understand the neural basis of episodic memory". *Journal of Neuroscience.* 27, 14365-74.

Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in Cognitive Science.* 11, 299–306.

Hassabis, D., & Maguire, E.A. (2009). The construction system of the brain. *Philosophical Transactions of the Royal Society of London. B: Biological Sciences.* 1521, 1263-1271.

Hoch, S. (1984). Counterfactual reasoning and accuracy in predicting personal events. *Journal of Experimental Psychology*. 11, 719-731.

Holmes, D.S. (1970). Differential change in affective intensity and the forgetting of unpleasant personal experiences. *Journal of Personality and Social Psychology.* 3, 234-239.

Klein, S. B., Loftus, J., & Kihlstrom, J. F. (2002). Memory and temporal experience: The effects of episodic memory loss on an amnesic patient's ability to remember the past and imagine the future. *Social Cognition*. 20, 353–379.

Kurczek, J.C., Duff, M.C., Jensen, U., Adolphs, R., Cohen, N.J., & Tranel, D. (2010). Hippocampal amnesia impairs the generation of mental representations across all time periods. *SfN Annual Meeting*, November 13–17, San Diego, CA. Abstr 600.8.

Martin, V.C., Schacter, D.L., Corballis, M, & Addis, D.R. (2011). A role for the hippocampus in encoding simulations of future events. *Proceedings of the National Academy of Science U.S.A*. 108, 13858-13863.

Meltzer, H. (1930). Individual differences in forgetting pleasant and unpleasant experiences. *Journal of Educational Psychology.* 21, 399- 409.

Okuda, J., Fujii, T., Ohtake, H., Tsukiura, T., Tanji, K., Suzuki, K., Kawashima, R., Fukuda, H., Itoh, M., & Yamadori, A. (2003). Thinking of the future and the past: The roles of the frontal pole and the medial temporal lobes. *Neuroimage*. 19, 1369-1380.

Race, E., Keane, M., & Verfaellie, M. (2011). Medial Temporal Lobe Damage Causes Deficits in Episodic Memory and Episodic Future Thinking Not Attributable to Deficits in Narrative Construction. *Journal of Neuroscience*. 31, 10262-10269.

Roese, N. (1997). Counterfactual thinking. *Psychological Bulletin*. 121: 133-148.

Rombouts, S.A.R.B., Scheltens, P., Machielsen, C.M., Barkhof, F., Hoogenraad, F.G.C., Veltmen, D.J., Valk, J., & Witter, M.P. (1999). Parametric fMRI analysis of visual encoding in the human medial temporal lobe. *Hippocampus*. 9, 637–643.

Schacter D.L., & Wagner A.D. (1999). Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. *Hippocampus.* 9, 7–24.

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society of London. B: Biological Sciences.* 362, 773–786.

Schacter, D. L., & Addis, D. R. (2007). The ghosts of past and future. *Nature*. *445*, 27.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). The prospective brain: Remembering the past to imagine the future. *Nature Reviews Neuroscience* 8, 657–661.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2008). Episodic simulation of future events: Concepts, data, and applications. *Annals of the New York Academy of Science.* 1124, 39–60.

Schacter, D.L., Addis, D.R., Hassabis, D., Martin, V.C., Spreng, R.N., & Szpunar, K.K. (2012). The future of memory: Remembering, imagining, and the brain. *Neuron, 76,* 677-694.

Sharot, T., Riccardi, A.M., Raio, C.M., & Phelps, E.A. (2007). Neural mechanisms mediating optimism bias. *Nature*. 450: 102-105.

Spreng, N., & Levine., B. (2006). The temporal distribution of past and future autobiographical events across the lifespan. *Memory and Cognition*. 34, 1644-1651.

Spreng R.N., Mar, R.A., & Kim, A.S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience*, 21: 489-510.

Squire, L.R., van der Horst, A., McDuff, S.G.R., Frascino, J.C., Hopkins, R.O., & Mauldin, K.N. (2010). Role of the hippocampus in remembering the past and imagining the future. *Proceedings of the National Academy of Science U.S.A*. 107: 19044-19048

Suddendorf, T., & Busby, J. (2005). Making decisions with the future in mind: Developmental and comparative identification of mental time travel. *Learning and Motivation.* 36, 110-125.

Szpunar, K. K., Watson, J. M., & McDermott, K. B. (2007). Neural substrates of envisioning the future. *Proceedings of the National Academy of Science U.S.A*. 104, 642–647.

Szpunar, K. K., & McDermott, K. B. (2008). Episodic future thought and its relation to remembering: Evidence from ratings of subjective experience. *Consciousness and Cognition.* 17, 330–334.

Szpunar, K.K., Chan, J.K.C., & McDermott, K.B. (2009). Contextual processing in episodic future thought. *Cerebral Cortex*. 19, 1539-1548.

Szpunar, K. K. (2010). Episodic future thought: An emerging concept. *Perspectives in Psychological Science.* 5: 142-162.

Szpunar, K.K., Addis, D. R., & Schacter, D.L. (2012). Memory for emotional simulations: Remembering a rosy future. *Psychological Science. 23*, 24-29.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist 25*, 1–12.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*. 53, 1–25.

Van Hoeck, N., Ma, N., Van Overwalle, F., & Vandekerckhove, M. (2010). Counterfactual thinking and the episodic system. *Behavioral Neurology*. 23, 225-227.

Van Hoeck, N., Ma, N., Ampe, L., Baetens, K., Vandekerckhove, M., & Van Overwalle, F. (in press). Counterfactual thinking: An fMRI study on changing the past for a better future. *Social Cognitive and Affective Neuroscience.* doi: 10.1093/scan/nss031

Waldfogel, S. (1948). The frequency and affective character of childhood memories. *Psychological Monographs.* 62, 291.

Walker, W. R., & Skowronski, J. J. (2009). The Fading Affect Bias….But what the hell is it for? *Applied Cognitive Psychology*. 23, 1122-1136.

Weiler, J.A., Suchan, B., & Daum, I. (2010). Foreseeing the future: Occurrence probability of imagined future events modulates hippocampal activation. *Hippocampus* 20, 685–690.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*. 39, 806-820.

Wilbur, W. R., Skowronski, J. J., & Thompson, C. P. (2003). Life is pleasant—and memory helps to keep it that way! *Review General Psychology*. 7, 203-210.

Williams, J. M., Ellis, N. C., Tyers, C., Healy, H., Rose, G., & MacLeod, A. K. (1996). The specificity of autobiographical memory and imageability of the future. *Memory and Cognition.* 24, 116–125.

Xue, G., Lu, Z., Levin, I.P. & Bechara, A. (2010). The impact of prior risk experiences on subsequent risky decision-making: The role of the insula. *Neuroimage*. 50:709-716.

Figure Legends

Figure 1. Experimental design. A screen with three episodic components (i.e., context, action/event, and outcome) and the specific condition as title (e.g., Positive) was presented to the participants for 10s. Participants were asked to press a button as soon as the construction period had finished (aproximately 4 to 5 seconds; see results), and were instructed to elaborate on their simulation for the remaing time. After a fixation cross, screens with the ratings of emotion and vividness, and then likelihood and regret/relief, were presented for 6 seconds each. Participants used a five-button MRI compatible response box to record their ratings. A final fixation cross indicated the end of the trial and beginning of next trial.

Figure 2. Latent variable 1 (LV 1). (A) Plot of brain scores with confidence intervals. Of note, confidence intervals are asymmetrical as a result of rescaling each condition's distribution during bootstraping. (B) Plot of temporal brain scores indicating weighted average of activation across all voxels in all subjects during the length of the task (TRs = 2 seconds) across all voxels in all subjects. (C) Brain regions associated with the experimental conditions at selected TRs. All regions are shown at a threshold of $p < .001$. (D) Plots indicating percent signal change of peak voxels in right parahippocampal gyrus (BA 28; xyz = 24 -22 -8), left posterior cingulate cortex (BA 23; xyz = -4 -32 22), and left superior parietal lobule (BA 7; xyz = -30 -70 52).

Figure 3. Latent variable 2 (LV 2). (A) Plot of brain scores with confidence intervals. Of note, confidence intervals are asymmetrical as a result of rescaling each condition's distribution during bootstraping. (B) Plot of temporal brain scores indicating weighted average of activation across all voxels in all subjects during the length of the task (TRs = 2 seconds) across all voxels in all subjects. (C) Brain regions associated with the experimental conditions at selected TRs. Brain regions associated with remember and likely counterfactuals are shown in yellow, while those associated with unlikely counterfactuals and the control condition are shown in cyan All regions are shown at a threshold of $p < .001$. (D) Plots indicating percent signal change of peak voxels in left parahippocampal gyrus (BA 19; xyz = -18 -48 -10), right parahippocampal gyrus (BA 19; xyz = 26 -52 -4), and left superior frontal gyrus (BA 10; xyz = -10 60 26).

Table 1: Regions associated with remember, positive and negative counterfactuals, versus the control condition (LV 1).

Region of Activation	Hemisphere	BA	MNI Coordinates			BSR*
			X	Y	Z	
TR 2 (2–4s after stimulus onset)						
Cerebellum	L		-18	-60	-20	-4.2757
Posterior Cingulate	L	23	-2	-38	20	-4.2746
Inferior Frontal Gyrus	R	9	44	-4	28	-4.9977
Cingulate Gyrus	R	24	24	-22	38	-4.8221
Parahippocampal Gyrus	R	28	24	-22	-8	-3.7748
TR 3 (4–6s after stimulus onset)						
Posterior Cingulate	L	23	-4	-32	22	-4.2725
Superior Temporal Gyrus	L	38	-34	20	-38	-3.7306
Precentral Gyrus	R	6	44	-8	26	-4.3638
Precuneus	R	7	16	-68	40	-4.5698
Cerebellum	R		20	-70	-46	-4.8541
TR 4 (6-8s after stimulus onset)						
Precuneus	L	7	-8	-64	48	-4.5656
Superior Parietal Lobule	L	7	-30	-70	52	-4.1429
Superior Temporal Gyrus	L - R	38	-40	20	-22	-4.0597
Precentral Gyrus	R	6	40	-10	38	-3.8172
TR 5 (8-10s after stimulus onset)						
Middle Temporal Gyrus	L	21	-52	-34	-6	-4.5411
Posterior Cingulate	L	29	-4	-42	14	-3.8372
TR 6 (10-12s after stimulus onset)						
Posterior Cingulate	L	23	-2	-52	22	-5.4876
Precuneus	L	7	-4	-68	34	-4.7842
Superior Parietal Lobule	L	7	-34	-70	44	-3.9137
Superior Temporal Gyrus	L	38	-38	14	-22	-4.2388

Note: All activations are reported for the first seven 2-second TRs. All activations reported survive a threshold of $p < .0002$ (BSR = 3.2). BA = approximate Brodmann area. L = Left; R = Right. * The bootstrap ratio (BSR) is the parameter estimate for that voxel over its standard error. It is proportional to a z score.

Table 2: Regions associated with remember and likely counterfactuals, versus unlikely counterfactuals and the control condition (LV 2).

			MNI Coordinates			
Region of Activation	Hemisphere	BA	X	Y	Z	BSR*
TR 2 (2–4s after stimulus onset)						
Superior Frontal Gyrus	L	10	-10	60	26	3.9903
TR 3 (4-6s after stimulus onset)						
Inferior Parietal Lobe	L - R	40	-40	-56	36	4.2571
Middle Frontal Gyrus	L	47	-42	36	-4	4.0301
Superior Frontal Gyrus	L	9	-12	52	30	3.924
Precuneus	R	7	20	-64	38	5.3604
Superior Temporal Gyrus	R	38	34	10	-38	5.0813
TR 4 (6-8s after stimulus onset)						
Superior Frontal Gyrus	L	9	-14	56	32	5.0309
Inferior Parietal Lobule	L	40	-38	-62	44	3.9532
Precuneus	R	7	4	-60	40	3.7858
Angular Gyrus	R	39	46	-70	36	3.7283
Superior Parietal Lobule	R	7	30	-60	44	3.6864
TR 5 (8-10s after stimulus onset)						
Superior Frontal Gyrus	L	9	-12	52	40	4.106
Inferior Frontal Gyrus	L	47	-60	24	-6	3.8246
Precuneus	L	7	-2	-60	36	4.3333
Inferior Parietal Lobule	L	40	-42	-60	38	3.7076
Middle Temporal Gyrus	L	21	-56	-4	-16	4.2505
TR 6 (10-12s after stimulus onset)						
Parahippocampal Gyrus	L	36	-24	-18	-30	4.948
Precuneus	L	7	-4	-60	40	4.9034
Inferior Parietal Lobule	L	40	-44	-62	40	3.8572
Inferior Temporal Gyrus	L	21	-58	-8	-14	5.1268
Middle Temporal Gyrus	L	39/22	-42	-64	30	4.0254

Note: All activations are reported for the first six 2-second TRs. All activations reported survive a threshold of $p < .0002$ (BSR = 3.2). BA = approximate Brodmann area. L = Left; R = Right. * The bootstrap ratio (BSR) is the parameter estimate for that voxel over its standard error. It is proportional to a z score.

Table 3: Regions associated with unlikely and the control condition, versus remember and likely counterfactuals (LV 2).

Region of Activation	Hemisphere	BA	MNI Coordinates			BSR*
			X	Y	Z	
TR 2 (2–4s after stimulus onset)						
Cingulate Gyrus	L	24	-14	4	38	-3.7572
Inferior Frontal Gyrus	R	47	28	30	4	-4.2578
Superior Frontal Gyrus	R	8/9	12	48	54	-4.1301
Middle Frontal Gyrus	R	46	52	30	28	-3.8343
Putamen	R		32	-14	8	-5.4074
Superior Temporal Gyrus	R	22	46	-26	-8	-5.5013
Middle Temporal Gyrus	R	21	62	0	-16	-3.8053
TR 3 (4-6s after stimulus onset)						
Superior Frontal Gyrus	L	6	-6	28	68	-3.8342
Middle Frontal Gyrus	L	10	-28	46	16	-3.7966
Parahippocampal Gyrus	L	30	-26	-54	2	-5.5633
Parahippocampal Gyrus	L	19	-18	-48	-10	-4.7548
Inferior Parietal Lobule	L	40	-60	-44	42	-5.8973
Claustrum	L - R		-30	14	8	-5.743
Insula	L	13	-58	-36	18	-5.0082
Caudate	L		-10	6	22	-4.2916
Middle Temporal Gyrus	L	22	-48	-38	6	-5.0794
Inferior Frontal Gyrus	R	45	60	14	20	-4.2282
Parahippocampal Gyrus	R	27	24	-36	-4	-5.9839
Postcentral Gyrus	R	3	34	-26	46	-4.9942
Thalamus	R		30	-32	6	-4.6507
Hypothalamus	R		6	-6	-12	-4.5517
TR 4 (6-8s after stimulus onset)						
Middle Frontal Gyrus	L	11	-26	48	-12	-5.3128
Precentral Gyrus	L - R	6	-38	-6	44	-4.3725
Anterior Cingulate	L - R	32	-14	22	26	-5.3403
Posterior Cingulate	L	29	-16	-44	6	-3.7754
Inferior Parietal Lobule	L	40	-58	-34	20	-3.6406
Putamen	L		-28	2	8	-4.4758
Medial Frontal Gyrus	R	25	10	22	-18	-4.4895
Inferior Frontal Gyrus	R	9	60	8	32	-4.0269
Parahippocampal Gyrus	R	19	26	-52	-4	-3.6186
Lingual Gyrus	R	19	24	-62	0	-3.9523
Thalamus	R		12	-38	8	-3.8167
Superior Temporal Gyrus	R	42/38	70	-20	6	-5.5658
TR 5 (8-10s after stimulus onset)						

Precentral Gyrus	L	6	-54	0	36	-5.1143
Cingulate Gyrus	L - R	32/31	-10	10	36	-4.4315
Middle Frontal Gyrus	L	8	-36	24	46	-3.6352
Postcentral Gyrus	L	4	-58	-32	20	-4.7878
Postcentral Gyrus	L - R	3	-22	-36	56	-4.5936
Middle Frontal Gyrus	R	10	44	48	18	-3.6474
Superior Temporal Gyrus	R	38	60	14	-10	-4.3752
Middle Temporal Gyrus	R	21	66	-28	-2	-3.6201
TR 6 (10-12s after stimulus onset)						
Middle Frontal Gyrus	L - R	10	-42	48	12	-4.1832
Superior Frontal Gyrus	L	9	-36	40	36	-3.9937
Cingulate Gyrus	L	31	-12	-36	42	-4.6973
Inferior Parietal Lobule	L	40	-60	-30	36	-4.944
Precuneus	L - R	7	-26	-50	52	-3.6877
Precentral Gyrus	R	44	52	8	10	-4.0911
Parahippocampal Gyrus	R	19	18	-54	-4	-5.0394
Middle Temporal Gyrus	R	20/21	50	-36	-10	-4.2747
Superior Temporal Gyrus	R	38	52	18	-16	-3.5878

Note: All activations are reported for the first six 2-second TRs. All activations reported survive a threshold of $p < .0002$ (BSR = 3.2). BA = approximate Brodmann area. L = Left; R = Right. * The bootstrap ratio (BSR) is the parameter estimate for that voxel over its standard error. It is proportional to a z score.

Table 4: Regions of activation revealed by the parametric modulation analysis.

Region of Activation	Hemisphere	BA	MNI Coordinates			t-value
			X	Y	Z	
A. Regions showing increasing activation modulated by subjective likelihood						
Positive Counterfactuals						
Inferior Frontal Gyrus	L	46	-44	38	4	3.44
Middle Frontal Gyrus	L	10	-36	58	8	3.36
Inferior Temporal Gyrus	L	20	-50	-24	-18	3.67
Negative Counterfactuals						
Precentral Gyrus	L - R	6	-46	-12	28	3.27
Anterior Cingulate	L	32/24	-6	40	8	3.42
Parahippocampal Gyrus	L	30	-24	-38	6	3.15
Insula	L	13	-40	-18	4	3.51
Thalamus	L - R		-22	-30	6	3.18
Caudate	L – R		-10	-6	20	2.87
Superior Temporal Gyrus	L	22	-58	-36	10	3
Superior Temporal Gyrus	L – R	38	-40	12	-28	2.84
Middle Frontal Gyrus	R	9/46	48	16	30	2.8
Inferior Frontal Gyrus	R	47	22	12	-16	2.79
Hippocampus	R		34	-10	-16	2.76
Peripheral Counterfactuals						
Posterior Cingulate	L	23	-2	-56	18	3.09
Parahippocampal Gyrus	L	19	-40	-52	-2	2.75
Middle Temporal Gyrus	L	19	-32	-60	14	3.47
Cuneus	L	17	-22	-78	0	3.34
Lingual Gyrus	L	18	-14	-80	-4	2.7
Insula	L - R	13	-34	-46	22	3.58
Middle Temporal Gyrus	L	39/21	-32	-54	22	3.47
Superior Temporal Gyrus	L - R	22	-62	-16	0	3.12
Middle Temporal Gyrus	L	37	-40	-62	-4	2.78
Caudate			32	-42	10	2.87
B. Regions showing decreasing activation modulated by subjective likelihood						
Positive Counterfactuals						
Cerebellum	L		-2	-44	-50	6.39
Caudate	L		-10	26	-6	4.22
Thalamus	L - R		-4	-16	6	3.76
Parahippocampal Gyrus	R	35	16	-22	-16	5.08
Superior Temporal Gyrus	R	42	2	-8	2	6.67
Negative Counterfactuals						
Middle Temporal/Angular Gyrus	L	39	-50	-72	30	4.11
Precuneus	R	7	6	-78	48	4.09

Peripheral Counterfactuals						
Postcentral Gyrus	L	1	-50	-28	56	4.19
Superior Frontal Gyrus	R	6	20	2	66	3.48
Cerebellum	R		2	-42	-12	4.82

Note: All activations reported survived a threshold of $p < .005$, $k = 10$. BA = approximate Brodmann Area. L = Left; R = Right.

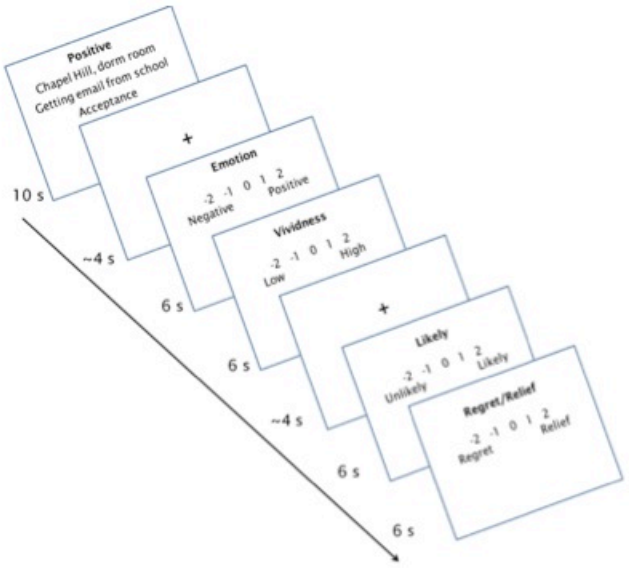
Positive
Chapel Hill, dorm room
Getting email from school
Acceptance
10 s
+
Emotion
-2 -1 0 1 2
Negative
Positive
~4 s
6 s
Vividness
-2 -1 0 1 2
Low
High
6 s
+
~4 s
Likely
-2 -1 0 1 2
Unlikely
Likely
6 s
Regret/Relief
-2 -1 0 1 2
Regret
Relief
6 s

Figure 1

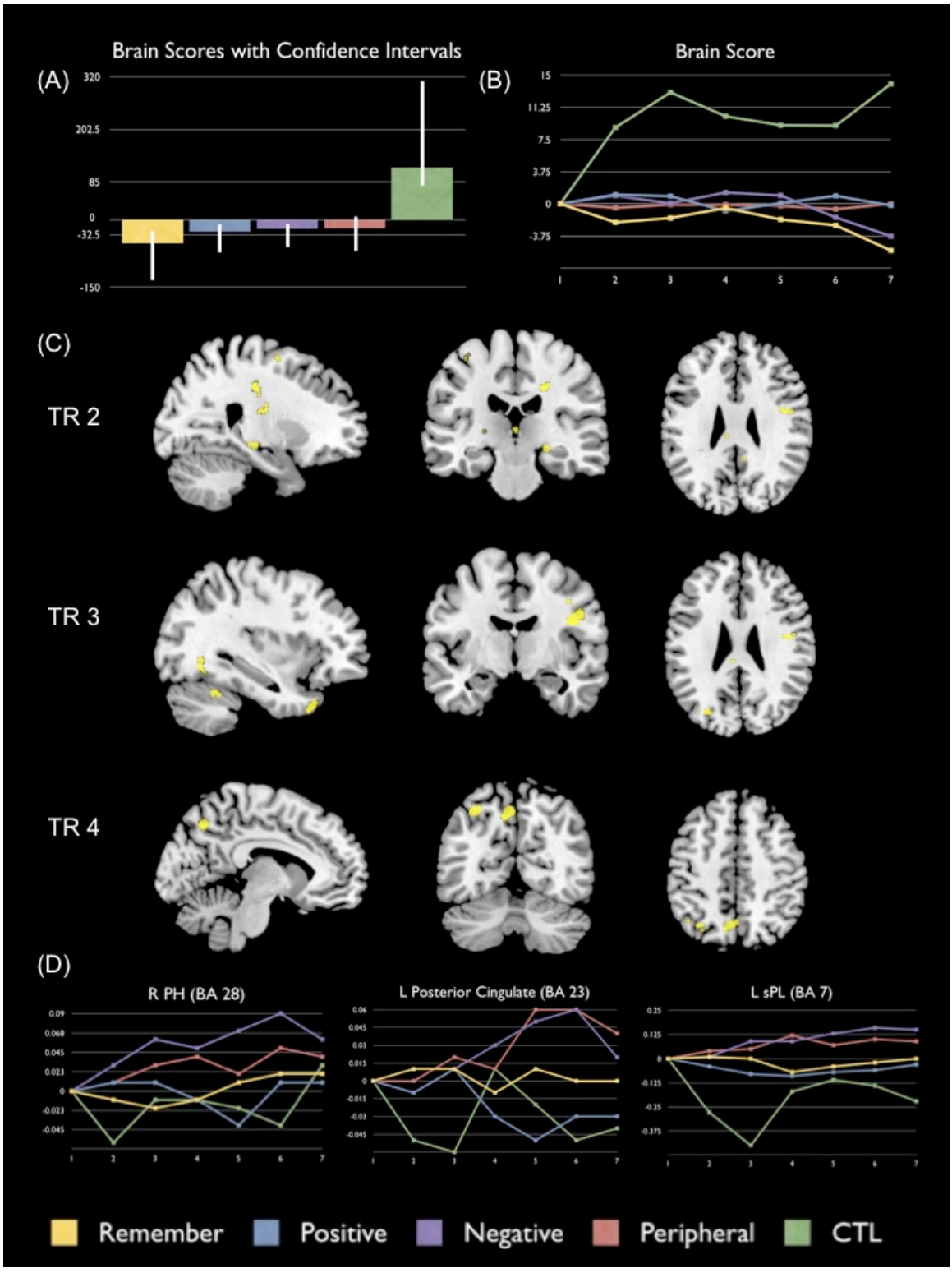
Brain Scores with Confidence Intervals
(A)
320
202.5
85
0
-32.5
-150
(B)
Brain Score
15
11.25
7.5
3.75
0
-3.75
1
2
3
4
5
6
7
(C)
TR 2
TR 3
TR 4
(D)
R PH (BA 28)
L Posterior Cingulate (BA 23)
L sPL (BA 7)
Remember
Positive
Negative
Peripheral
CTL

Figure 2

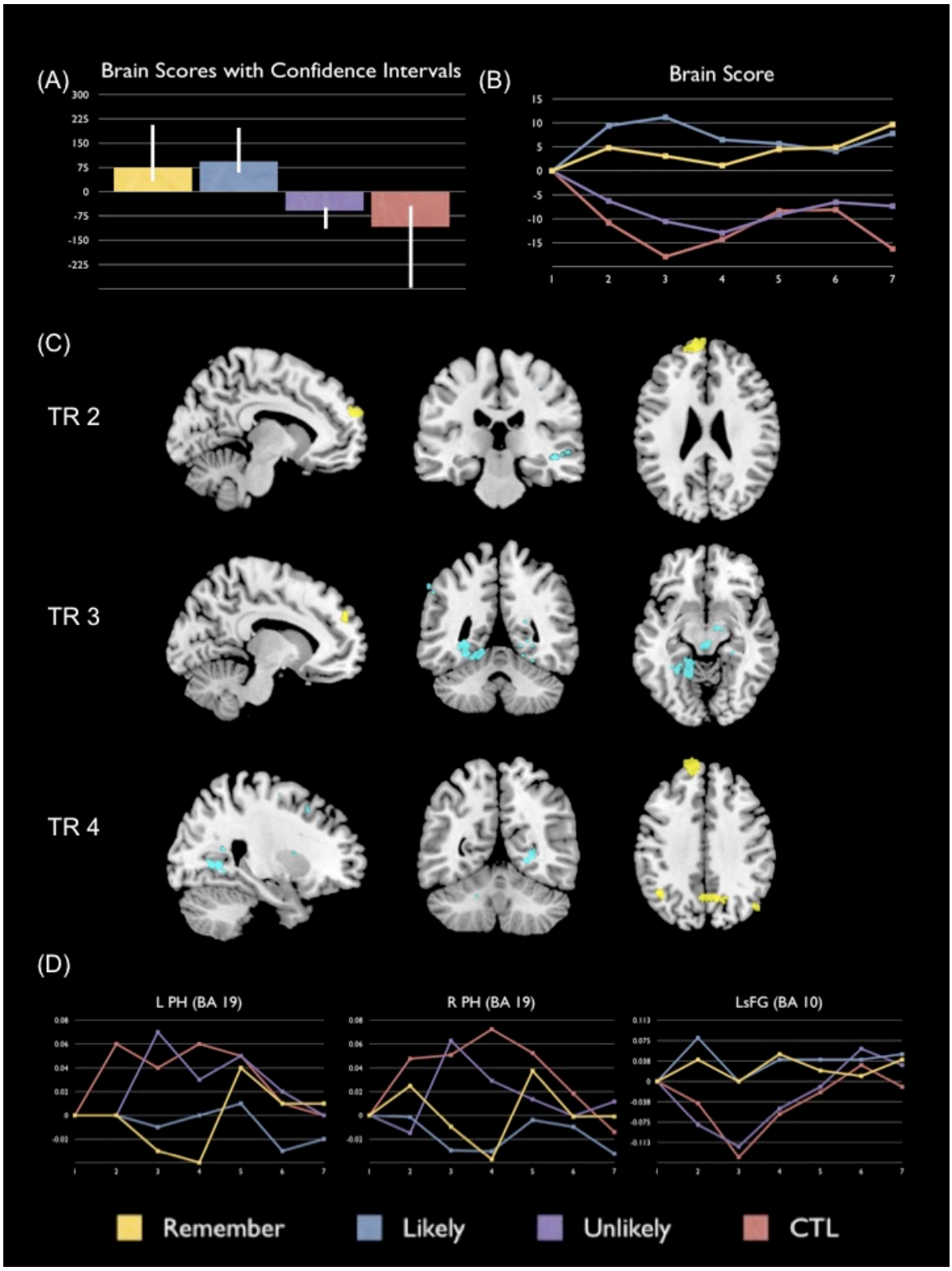
(A)
Brain Scores with Confidence Intervals
300
225
150
75
0
-75
-150
-225
(B)
Brain Score
15
10
5
0
-5
-10
-15
1
2
3
4
5
6
7
(C)
TR 2
TR 3
TR 4
(D)
L PH (BA 19)
R PH (BA 19)
LsFG (BA 10)
Remember
Likely
Unlikely
CTL

Figure 3

Table 5 (Supplementary material): Regions associated with remember, likely, and unlikely counterfactuals versus the control condition (LV 3).

Region of Activation	Hemisphere	BA	MNI Coordinates			BSR*
			X	Y	Z	
TR 2 (2–4s after stimulus onset)						
Posterior Cingulate	L	23	-2	-34	20	4.4041
Thalamus	L		0	-20	6	4.9773
Medial Frontal/Precentral gyrus	R	6	6	-10	54	4.4198
Parahippocampal Gyrus	R	28	24	-20	-10	3.6014
Caudate	R		22	-14	22	3.5777
TR 3 (4-6s after stimulus onset)						
Posterior Cingulate	L	23	0	-32	22	4.4206
Precuneus	L - R	7	0	-64	40	3.7845
Middle Temporal Gyrus	L	37	-42	-56	-6	5.4421
Middle Temporal/Temporal Pole	L - R	21/22	-56	2	-10	3.8429
Precentral Gyrus	R	6	44	-8	26	4.1039
Cingulate gyrus	R	24	4	-14	44	4.5066
Hippocampus	R		34	-28	-10	3.7142
Thalamus	R		22	-28	10	5.1408
TR 4 (6-8s after stimulus onset)						
Middle Temporal Gyrus/Precuneus	L	7	-4	-64	50	3.9015
Middle Temporal/Temporal Pole	L	21	-56	-30	-2	4.0394
TR 5 (8-10s after stimulus onset)						
Inferior Frontal Gyrus	L	47	-56	32	-12	3.9543
Posterior Cingulate	L	29	-4	-40	16	5.1475
Precuneus	L	7	-2	-78	34	4.1333
Middle /Inferior Temporal Gyrus	L	21	-52	-34	-6	4.4684
TR 6 (10-12s after stimulus onset)						
Posterior Cingulate	L	29	0	-42	18	3.6953
Precuneus	L - R	7	0	-66	34	5.3233
Inferior Parietal Lobule	L	40	-46	-58	34	4.5137

Note: All activations are reported for the first six 2-second TRs. All activations reported survive a threshold of $p < .0002$ (BSR = 3.2). BA = approximate Brodmann area. L = Left; R = Right. * The bootstrap ratio (BSR) is the parameter estimate for that voxel over its standard error. It is proportional to a z score.

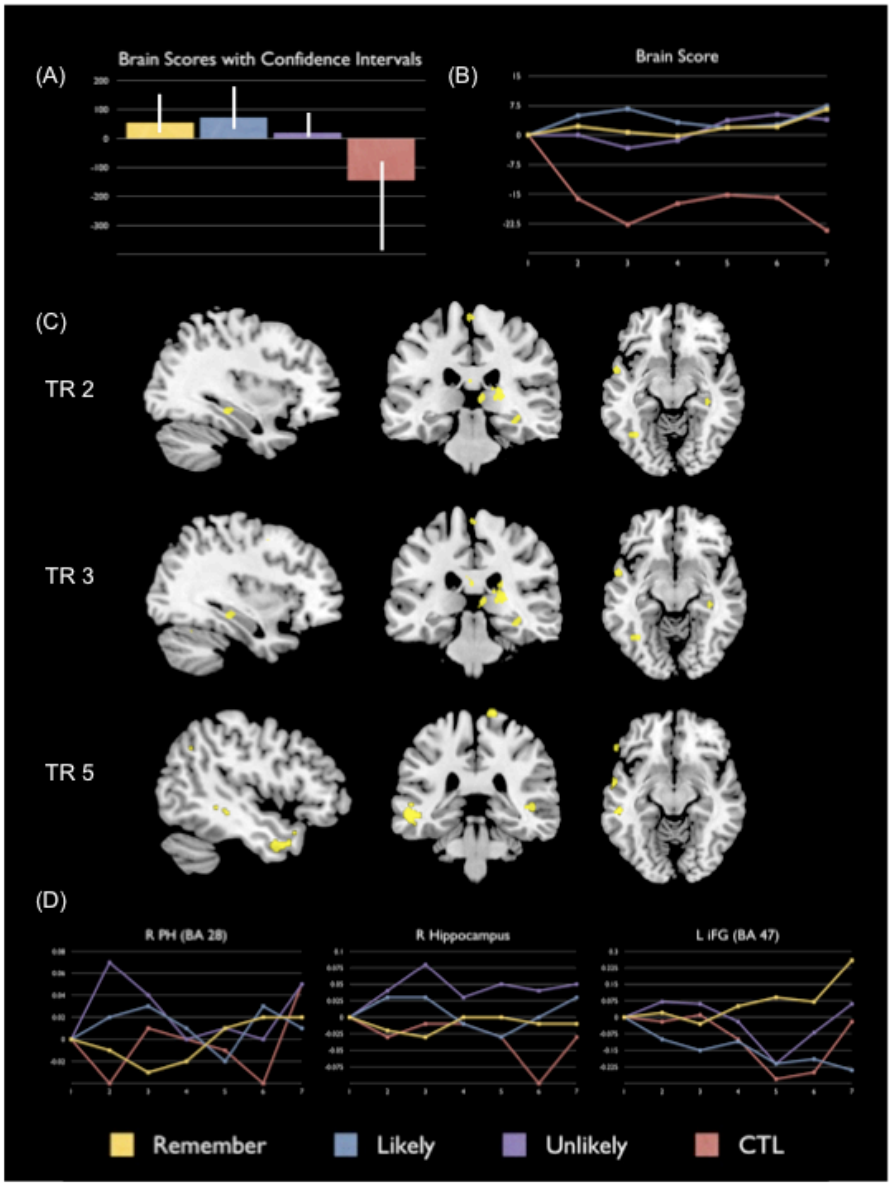

Figure 4. Latent variable 3 (LV 3). (A) Plot of brain scores with confidence intervals. Of note, confidence intervals are asymmetrical as a result of rescaling each condition's distribution during bootstraping. (B) Plot of temporal brain scores indicating weighted average of activation across all voxels in all subjects during the length of the task (TRs = 2 seconds) across all voxels in all subjects. (C) Brain regions associated with the experimental conditions at selected TRs. Brain regions associated with remember and likely counterfactuals are shown in yellow; brain regions associated with unlikely counterfactuals and the control condition are shown in cyan All regions are shown at a threshold of p <.001. (D) Plots indicating percent signal change of peak voxels in right parahippocampal gyrus (BA 28; xyz = 24 -20 -10), right hippocampus (xyz = 34 -28 -10), and left inferior frontal gyrus (BA 47; xyz = -56 32 -12).

Table 6. (Supplementary material). Regions associated with remembering and likely counterfactuals versus unlikely counterfactuals (LV 4)

Region of Activation	Hemisphere	BA	MNI Coordinates			BSR*
			X	Y	Z	
TR 2 (2–4s after stimulus onset)						
Superior Frontal Gyrus	L	10	-10	60	28	5.191
Inferior Frontal Gyrus +	L	45	-54	20	10	3.972
Superior Temporal Gyrus +	L	22	-54	16	-2	3.76
Middle Temporal Gyrus +	L	19	-38	-82	22	4.911
Middle Frontal Gyrus +	R	8/6	40	2	64	6.607
TR 3 (4-6s after stimulus onset)						
Middle Frontal Gyrus	L	47	-44	34	-4	4.348
Middle Temporal Gyrus +	L	19	-40	-80	20	4.211
Inferior Parietal Lobule	L - R	40	-42	-54	-40	4.646
TR 4 (6-8s after stimulus onset)						
Superior Frontal Gyrus	L	9	-18	62	36	3.689
Inferior Frontal Gyrus +	L	46/47	-48	30	14	4.351
Inferior Parietal Lobule	L	40	-40	-58	38	4.064
Cerebellum +	L		-22	-70	-42	4.212
Parahippocampal Gyrus +	R	34	12	-12	-28	3.793
TR 5 (8-10s after stimulus onset)						
Superior Frontal Gyrus	L	9	-16	54	32	4.615
Inferior Frontal Gyrus	L	46	-46	26	16	3.736
TR 6 (10-12s after stimulus onset)						
Medial Frontal Gyrus +	L	8	-16	32	50	4.2429
Inferior Temporal Gyrus	L	21	-58	-10	-16	6.206
Parahippocampal Gyrus	L	35	-26	-12	-28	6.134
Caudate +	L		-12	24	14	3.708
Cerebellum +	L		-16	-36	-36	4.584
Medial Frontal Gyrus +	R	10	8	66	4	3.880

Note: All activations are reported for the first six 2-second TRs. All activations reported survive a threshold of $p < .0002$ (BSR = 3.2). BA = approximate Brodmann area. L = Left; R = Right. * The bootstrap ratio (BSR) is the parameter estimate for that voxel over its standard error. It is proportional to a z score. + Indicates regions that were not identified in LV 2 but showed activation in LV 4.

Table 7. (Supplementary material). Regions associated with unlikely counterfactuals versus remembering and likely counterfactuals (LV 4).

Region of Activation	Hemisphere	BA	MNI Coordinates			BSR*
			X	Y	Z	
TR 2 (2–4s after stimulus onset)						
Superior Frontal Gyrus	L	8	-6	44	56	-3.741
Superior/Middle Temporal Gyrus	L – R	21	-62	-32	-16	-3.739
Cingulate Gyrus	L	24	-6	-4	42	-5.3796
Inferior Frontal Gyrus	R	45	62	26	6	-3.772
Putamen	R	38	32	-14	8	-4.219
Fusiform Gyrus +	R	37	32	-38	-18	-4.834
Superior Temporal Gyrus +	R	38	48	16	-36	-3.872
Insula +	R	40	50	-24	14	-5.302
Thalamus +	R		20	-34	0	-7.358
TR 3 (4-6s after stimulus onset)						
Inferior Parietal Lobule	L	40	-60	-42	38	-5.203
Parahippocampal Gyrus	L - R	30	-28	-52	0	-5.210
Superior Temporal Gyrus	L	41/42	-46	-40	8	-6.581
Insula	L - R	13	-28	-40	26	-4.083
Claustrum	L		-32	10	6	-4.914
Cerebellum	L - R		-14	-44	-8	-5.205
Fusiform Gyrus +	R	37	48	-36	-12	-4.013
Postcentral Gyrus	R	2	38	-28	46	-6.057
Superior Temporal Gyrus	R	22	68	-20	2	-4.421
TR 4 (6-8s after stimulus onset)						
Middle Frontal Gyrus	L	9	-34	26	34	-4.26
Precentral Gyrus	L	6	-24	-16	58	-4.725
Inferior Parietal Lobule +	L	40	-40	-32	44	-4.324
Postcentral Gyrus	L	40	-58	-30	20	-4.767
Parahippocampal Gyrus +	L	30	-28	-58	4	-4.861
Putamen	L		-30	-2	8	-4.230
Cerebellum +	L - R		0	-54	-42	-3.795
Inferior Frontal Gyrus	R	44	62	8	20	-4.145
Postcentral Gyrus +	R	3	36	-30	50	-4.033
Cingulate Gyrus	R	24/32	10	12	44	-6.320
Parahippocampal Gyrus	R	18	28	-60	2	-5.583
Superior Temporal Gyrus	R	22/38	64	-22	4	-5.023
TR 5 (8-10s after stimulus onset)						
Middle Frontal Gyrus	L	6	-26	-10	54	3.983
Precentral Gyrus	L	6	-40	-6	42	-4.556
Postcentral Gyrus	L - R	2	-52	-30	44	-5.250

Postcentral Gyrus	L	3	-58	-18	30	-4.849
Cingulate Gyrus	L - R	24	-4	14	32	-5.233
Inferior Parietal Lobule +	L	40	-38	-36	44	-4.127
Superior Temporal Gyrus	R	21/22	66	-24	-2	-4.619
Cerebellum +	R		22	-38	-14	-4.359
TR 6 (10-12s after stimulus onset)						
Middle Frontal Gyrus	L - R	10	-40	48	14	-3.698
Inferior Parietal Lobule	L	40	-58	-32	24	-4.205
Postcentral Gyrus	R	7	8	-54	76	-4.018
Cingulate Gyrus	R	32	4	10	44	-4.286
Superior Temporal Gyrus	R	41	56	-28	12	-4.205
Lingual Gyrus	R	19	16	-52	-4	-6.116
Hippocampus +	R		36	-22	-10	-4.065
Thalamus +	R		18	-36	0	-4.404

Note: All activations are reported for the first six 2-second TRs. All activations reported survive a threshold of $p < .0002$ (BSR = 3.2). BA = approximate Brodmann area. L = Left; R = Right. * The bootstrap ratio (BSR) is the parameter estimate for that voxel over its standard error. It is proportional to a z score. + Indicates regions that were not identified in LV 2 but showed activation in LV 4.

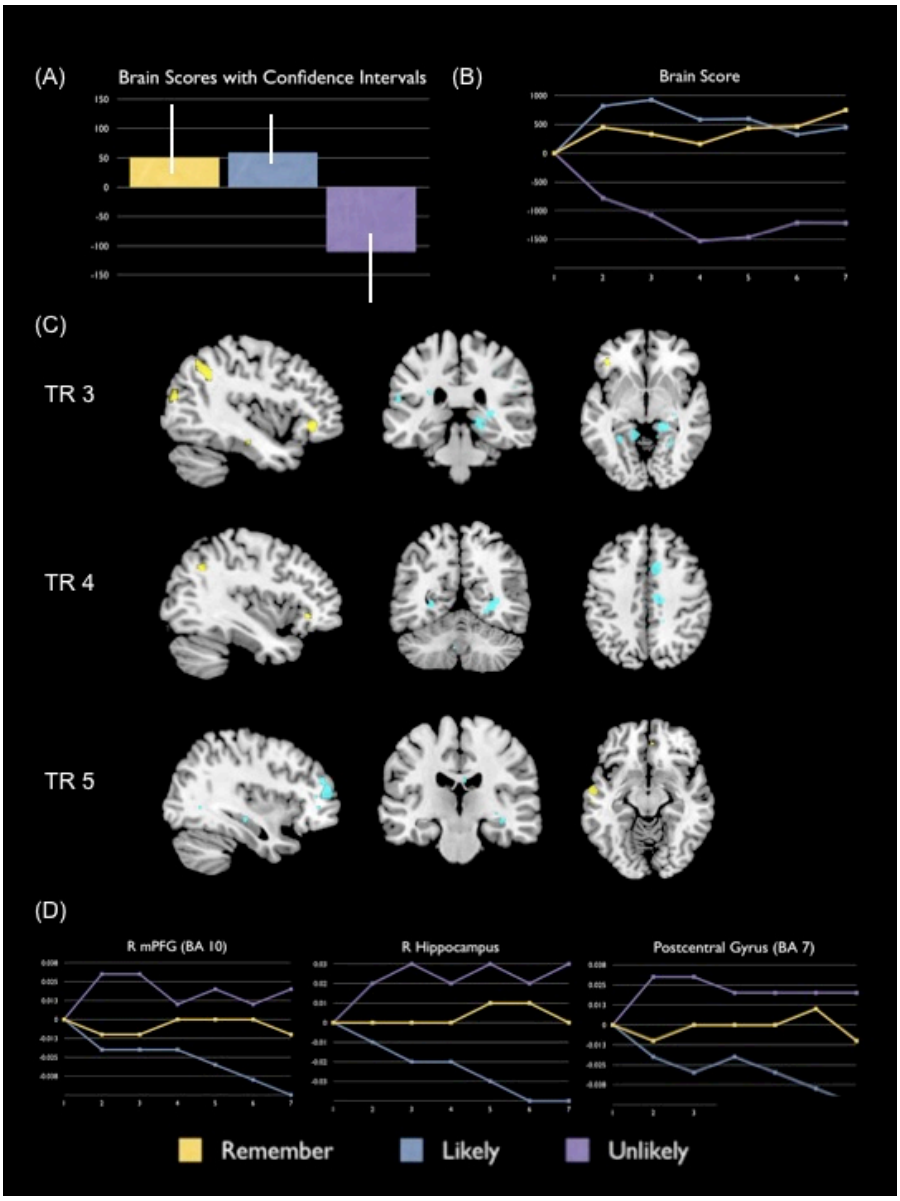

Figure 5. Latent variable 3 (LV 4). (A) Plot of brain scores with confidence intervals. Of note, confidence intervals are asymmetrical as a result of rescaling each condition's distribution during bootstraping. (B) Plot of temporal brain scores indicating weighted average of activation across all voxels in all subjects during the length of the task (TRs = 2 seconds) across all voxels in all subjects. (C) Brain regions associated with the experimental conditions at selected TRs. Brain regions associated with remember and likely counterfactuals are shown in yellow; brain regions associatied with unlikely counterfactuals are shown in cyan All regions are shown at a threshold of p <.001. (D) Plots indicating percental signal change of peak voxels in right middle frontal gyrus (BA10; xyz = 8 66 4), right hippocampus (xyz = 36 -22 -10) and right postcentral gyrus (BA 7; xyz = 8 -54 76).